The relative correspondence of cranial and genetic distances in papionin taxa and the impact of allometric adjustments

Heather F. Smith[a,b*], Noreen von Cramon-Taubadel[c]

[a] Department of Anatomy, Arizona College of Osteopathic Medicine, Midwestern University.
[b] School of Human Evolution and Social Change, Arizona State University
[c] Department of Anthropology, University at Buffalo, SUNY, Buffalo, NY
*Corresponding author: Heather F. Smith, Department of Anatomy, Arizona College of Osteopathic Medicine, Midwestern University, 19555 N. 59th Ave., Glendale AZ USA.
E-mail: hsmith@midwestern.edu, Tel: 1-623-572-3726.

Text pages: 27
Tables: 10
Figures: 5

Keywords: papionini, geometric morphometrics, cranial regions, genetic distance

Running header: Cranial and genetic distance correlates in papionins

Abstract

The reconstruction of phylogenetic relationships in the primate fossil record is dependent upon a thorough understanding of the phylogenetic utility of craniodental characters. Here, we test three previously proposed hypotheses for the propensity of primate craniomandibular data to exhibit homoplasy using a study design based on the relative congruence between cranial distance matrices and a consensus genetic distance matrix ("genetic congruence") for papionin taxa: 1. Matrices based on cranial regions subjected to less masticatory strain are more genetically congruent than high-strain cranial matrices; 2. Matrices based on cranial regions developing earlier in ontogeny are more genetically congruent than matrices based on regions that develop later; 3. Matrices based on cranial regions with greater anatomical/functional complexity are more genetically congruent than matrices based on anatomically simpler regions.

Morphological distance matrices based on the shape of 15 different cranial regions, delineated on the basis of previous catarrhine studies, were statistically compared to matrices of known genetic distances in papionins. Since sexual dimorphism and allometry are known to characterize this clade, several analytical iterations were conducted: 1) mixed-sex, male-only, and female-only analyses; 2) with and without an allometric scaling adjustment. Across all data sets, the chondrocranium matrix was the most consistently correlated with genetic distances, which is also consistent with previous studies of cercopithecoid taxa, however there was no consistent support for the internal predictions of the three hypotheses tested. Allometric scaling corrections had the largest impact on the genetic congruence of facial shape matrices, a result which is consistent with previous studies that have described facial homoplasy in papionin taxa. These findings differ from patterns described for hominoid taxa, suggesting that no single predictive criterion can explain phylogenetic utility of cranial datasets across catarrhine primate taxa. Many of the differences in morphological-genetic matrix correlations could result from different levels of phenotypic integration and evolvability in cercopithecoids and hominoids, suggesting that further study of these phenomena in extant primates is warranted.

Introduction

Accurate reconstructions and interpretations of the past, including the taxonomy, phylogeny, and evolutionary adaptations of fossil taxa require a thorough understanding of the biology of extant taxa. In particular, inferences regarding the hominin and non-hominin primate fossil records are dependent upon the development of an accurate inference model of primate morphological diversity. The vast majority of the primate fossil record is composed of specimens for which direct DNA evidence is unattainable. Thus, morphological data must necessarily form the central basis for many phylogenetic, systematic, and evolutionary retrodictions. In primates, there is a general consensus that cranial morphology reflects genetic relationships among species and individuals to a reasonable degree; however, documented instances of homoplasy and phenotypic plasticity complicate such interpretations, and result in potentially contradictory information among morphological datasets.

Most notably, Collard and Wood (2000) determined that several sets of craniomandibular and dental characters in hominoid and papionin species did not reflect the known phylogenetic relationships among these taxa, suggesting that craniodental characters may not be reliable indicators of phylogeny in the fossil record. Several subsequent studies have also revealed that different subsets of cranial data differ in their correspondence with genetic distances in human populations (Roseman, 2004; Harvati and Weaver, 2006; Smith, 2009; von Cramon-Taubadel, 2009a; b; 2011b; a), guenons (Cardini and Elton 2008) and hominoid taxa (von Cramon-Taubadel and Smith, 2012). These findings highlight the importance of understanding the evolutionary basis for patterns of cranial variation across extant primate clades before they are extrapolated into the fossil record.

In the paleo(anthropological) literature, three major factors, thought to predict the phylogenetic utility of different craniodental datasets, have been suggested and widely discussed (for review see (von Cramon-Taubadel, 2014). These factors comprise (1) the extent to which particular aspects of cranial morphology might be influenced by biomechanical (typically masticatory) strain, (2) the ontogenetic development and perceived “heritability” of cranial regions, and (3) the overall anatomical and functional complexity of particular cranial elements.

Below we detail how these factors have been employed as a theoretical basis for proposed phylogenetic hypotheses, and the outcomes of previous tests of these hypotheses.

Biomechanical strain

The "homoiology hypothesis" is predicated on the idea that skeletal regions that respond to non-genetic factors such as biomechanical stress (i.e. are subject to phenotypic plasticity) are more likely to result in homoplasies ("homoiologies") than bones under reduced loading regimes (Collard and O'Higgins, 2001; Wood and Lieberman, 2001; Lycett and Collard, 2005; Collard and Wood, 2007). In other words, bones from higher strain regions should be more variable morphologically, and consequently are thought to be less reliable indicators of phylogeny, than those from lower strain regions (Lieberman, 1995). In the skull, the primary biomechanical stressors derive from the masticatory apparatus, including the strains experienced by the craniofacial bones during the chewing cycle, the forces resulting from muscle vectors in the temporalis, masseter, and pterygoid muscles, and the stress experienced via the load-bearing temporomandibular joint (Bouvier, 1986; Wall, 1999; Vinyard et al., 2003).

Wood and Lieberman (2001) compared levels of morphological variation in lower strain cranial regions (basicranium, neurocranium, upper face) to those that experience higher levels of masticatory strain (palate, mandible) in several catarrhine primates, and determined that on average, the higher strain regions were indeed more variable than lower strain regions. From this finding, they concluded that highly plastic characters, such as those of the masticatory complex, should be avoided in taxonomic evaluations of fossil hominin specimens due to their inherent variability and presumed unreliability (Wood and Lieberman, 2001). However, subsequent analyses directly comparing phylogenetic trees based on craniodental data and molecular genetic data found that, despite elevated levels of variability, high strain cranial regions did not produce trees any less congruent with the molecular phylogeny than low strain regions in papionins (Lycett and Collard, 2005) or hominoids (Collard and Wood, 2007; von Cramon-Taubadel and Smith, 2012). In fact, in contrast to the homoiology hypothesis, several of the more phenotypically plastic regions were found to reflect the genetic relationships more reliably than those that remodel less in response to environmental stimuli (Collard and Wood, 2007).

Recently, the homoiology hypothesis was tested in hominoids using 3D landmark-based data and a morphological-genetic distance matrix correlation approach (von Cramon-Taubadel and Smith, 2012). Morphological matrices based on the shape of the mandible and palatomaxilla matrices were found to be less strongly correlated with the genetic matrix than many of the other cranial regions; however, they still yielded morphological matrices that were significantly correlated with the genetic distances among taxa, albeit with lower correlation coefficients. Interestingly, other "masticatory" regions, such as the zygotemporal were among the most strongly correlated cranial matrices with the genetic distance matrix (von Cramon-Taubadel and Smith, 2012). In guenons, the morphology of several masticatory regions, the mandible, oral cavity, and zygomatic region, all yielded morphological matrices significantly correlated with genetic distances among taxa, but the non-masticatory chondrocranium did so with a higher correlation coefficient (Cardini and Elton, 2008). Therefore, despite the intuitive link between masticatory strain, plasticity, and homoplasy, the relationship between these factors is complex, and this phenomenon does not necessarily negatively affect the correspondence between cranial morphology and genetic patterns (von Cramon-Taubadel, 2009b; Roseman et al., 2010).

The homoiology hypothesis has also been tested empirically in humans (von Cramon-Taubadel, 2009b). A comparison of cranial regions associated with masticatory function (palatomaxilla, zygotemporal) versus those that are unassociated with mastication (basicranium, neurocranium, upper face) in 12 human populations revealed that "masticatory" cranial regions are more variable in humans, as they are in other catarrhine species (von Cramon-Taubadel, 2009b). However, this increased phenotypic plasticity does not impact the extent to which these regions reflect past population history, because some masticatory and non-masticatory cranial regions were found to be equally genetically congruent when morphological and neutral genetic among-population distance matrices were compared (von Cramon-Taubadel, 2009b).

In humans, a number of studies have indicated that the morphology of the masticatory apparatus may be influenced by subsistence strategies (Larsen, 1997; González-José et al., 2005; Sardi et al., 2006; Lieberman, 2008; Pinhasi et al., 2008; Paschetta et al., 2010; Holmes and Ruff, 2011; Lieberman, 2011; von Cramon-Taubadel, 2011b). In particular, morphology of the mandible shows a weak correspondence with geographic distance (Nicholson and Harvati, 2006), and a

lower correlation with neutral genetic distances than many other cranial regions (Smith, 2009). A comparison of the morphology of masticatory regions, such as the mandible and palatomaxilla, revealed that they co-vary with subsistence behavior among human populations (von Cramon-Taubadel, 2011a). Thus, it appears that the masticatory regions of the craniofacial complex in humans reflect a myriad of neutral and non-neutral evolutionary factors, thus influencing their covariation with inter-population genetic relationships.

Development and heritability

Several authors have also suggested a direct link between trait heritability, phenotypic plasticity, and developmental variation, such that presumably plastic and, therefore, variable traits are assumed to also exhibit lower heritability (Lieberman et al., 1996; Wood and Lieberman, 2001). The rationale behind this argument is similar to that of the homoiology hypothesis in that morphological regions thought not to be affected by phenotypic plasticity are predicted to more accurately reflect underlying "genetic" factors, rather than environmentally-induced sources of variation during an organism's lifetime.

In particular, it is often suggested that of the three major developmental complexes of the cranium—splanchnocranium, neurocranium, and basicranium—the endochondrally ossifying basicranium should be most developmentally stable (Olson, 1981; Lieberman et al., 1996; Strait, 2001; Wood and Lieberman, 2001) and therefore, offer a more reliable estimate of underlying inherited information. The basicranium develops upon a stable cartilaginous template and completes its ossification earlier in ontogeny than the other two developmentally-defined cranial regions. Thus, the basicranium is expected to be less susceptible to homoplastic changes than the facial skeleton and neurocranium, because it is considered to be more heritable and buffered from environmental disturbances during ontogeny (Olson, 1981; Strait et al., 1997; Lieberman et al., 2000a; Lieberman et al., 2000b). As a result, the basicranium has formed the basis for many comparative and phylogenetic studies of fossil hominin and non-hominin primates (Lieberman et al., 1996; Strait, 2001; Nevell and Wood, 2008).

The hypothesis that the basicranium, with its presumed higher heritability, should yield matrices that more strongly correlate with genetic distance matrices than other cranial regions,

such as the splanchnocranium (face) and neurocranium, has been explicitly tested within humans (von Cramon-Taubadel, 2011b) and across hominoid taxa (von Cramon-Taubadel and Smith, 2012). Among human populations, matrices based on basicranial shape have been shown to yield significant correlations with genetic population distances (Smith, 2009; von Cramon-Taubadel, 2011b); however, a comparison with other cranial regions revealed that it is not significantly more genetically congruent than the splanchnocranium or neurocranium (von Cramon-Taubadel, 2011b). These findings suggest that despite the strong correlation between basicranial shape and genetic relationships in humans, developmental factors alone cannot explain this association. In hominoid taxa, the basicranium was found to be among the least genetically congruent of the cranial regions, although it still produces a morphological matrix that is significantly correlated with genetic distances (von Cramon-Taubadel and Smith, 2012).

The phylogenetic utility of basicranial shape has also been tested in several cercopithecoid species (Cardini and Elton, 2008; Gilbert, 2011). Among guenons, the basicranium was found to be the cranial region that produced phenotypic patterns most highly correlated with the published consensus phylogeny. However, differences in basicranial morphology among taxa did not contain the degree of resolution necessary to reconstruct the precise topology of the consensus phylogenetic tree (Cardini and Elton, 2008). In addition, the face, dermatocranium, and neurocranium all yielded matrices that were also significantly correlated with the genetic distance matrix (Cardini and Elton, 2008). A study of the correspondence between basicranial morphology and the consensus phylogeny in another cercopithecoid group, the papionins, found that 3D basicranial morphology corrected for size-correlated information yielded phenetic trees that are inconsistent with the consensus genetic topology for various papionin genera (Gilbert, 2011). Thus, despite the intuitive association between developmental stability, heritability, and phylogenetic utility, there seems to be little consistent support for the "basicranial hypothesis" across catarrhine primates.

Recently, a quantitative genetic study of captive baboons found no significant differences in heritability across traits from different cranial regions (Roseman et al., 2010), suggesting that there is no *a priori* reason to suspect that cranial regions will perform differently in terms of their phylogenetic efficacy. Moreover, and in contrast with previous studies of wild papionin taxa (e.g.,

Lycett and Collard, 2005), they find no disparity in intrinsic phenotypic variation among cranial traits that could be attributable to differences in either additive genetic or environmental variation. Nevertheless, empirical studies of wild primates from across a range of taxa have demonstrated differences in terms of both phenotypic variability and phylogenetic efficacy of different craniomandibular and dental traits (see von Cramon-Taubadel, 2014 for review), even if heritability, phenotypic variability, and phylogenetic efficacy are not evidently causally related (von Cramon-Taubadel, 2009b; Roseman et al., 2010).

Anatomical and functional complexity

The third major rationale that has been employed to explain why some cranial regions might reflect genetic relationships more accurately than others centers on the concept of anatomical complexity (Lockwood et al., 2004). The logic behind this idea is that, for aspects of anatomy that are functionally and/or structurally complex, no single adaptive force can produce sizeable homoplastic differences among distantly related species. In other words, these complex regions are relatively immune to the effects of diversifying selection. This pattern is thought to be especially likely to characterize anatomical complexes that participate in multiple functions that are crucial for survival, such as complex locomotor or sensory functions (Lockwood et al., 2004).

Lockwood and colleagues (2004) invoked the anatomical complexity hypothesis as an explanation for the phylogenetic utility of the temporal bone. They argued that the intricate three-dimensional morphology of the temporal bone, coupled with its roles in locomotion, equilibrium, audition, and mastication, rendered it minimally homoplastic, and therefore phylogenetically informative (Lockwood et al., 2004). They also noted that temporal bone form is influenced by encephalization and concomitant cognition; functional factors which are thought to contribute to high levels of anatomical constraint (Lockwood et al., 2004).

Lockwood and colleagues (2004) were the first to demonstrate that the 3D morphology of the temporal bone could be used to reconstruct the consensus phylogeny of hominoid primates, and many subsequent studies have supported this assertion (Harvati and Weaver, 2006; Smith et al., 2007; Smith, 2009; von Cramon-Taubadel, 2009a; von Cramon-Taubadel and Smith, 2012). However, it has been shown that, in the case of tree-based phylogenetic analyses, the choice of

outgroup - in this case *Hylobates* versus *Pongo* - may dictate whether the correct topology is recovered (Bjarnason et al., 2011). In a comparison of morphological distance matrices based on numerous cranial regions to published genetic distances among hominoids, von Cramon-Taubadel and Smith (2012) found that temporal bone shape produces morphological matrices that are correlated with genetic distances among taxa, but that they were not significantly more strongly correlated than matrices based on other, less anatomically complex, cranial bones. In fact, the frontal bone yielded morphological distances matrices that were more strongly correlated with the genetic distance matrix than the temporal bone (von Cramon-Taubadel and Smith, 2012).

Several studies of modern humans have demonstrated that among-population distance matrices based on temporal bone morphology are significantly correlated with genetic distances among populations (Harvati and Weaver, 2006; Smith et al., 2007; Smith, 2009; von Cramon-Taubadel, 2009a; Smith et al., 2013). Comparisons of the extent to which a morphological distances matrix based on the temporal bone is correlated with neutral genetic distances relative to other cranial bone matrices revealed that the temporal bone is significantly more genetically congruent than the maxilla, occipital, and zygomatic bones (von Cramon-Taubadel, 2009a). However, when its morphology is considered in relation to other relatively complex regions of the skull, such as the neurocranium, upper face, and entire chondrocranium, the temporal bone is not any more genetically congruent than these other regions (Smith, 2009).

Cranial morphology and allometry in papionin primates

Previous studies of cranial morphology in the papionin taxa have suggested relatively high levels of craniomorphic homoplasy in this tribe (Fleagle and McGraw, 1999; Lockwood and Fleagle, 1999; McGraw and Fleagle, 2006; Gilbert, 2007; Gilbert et al., 2009). Phylogenetic studies based on traditional morphometric variables resulted in three primary groups of papionins initially being recognized: mangabeys, baboons, and geladas (Thorington and Groves, 1970; Jolly, 1972; Szalay and Delson, 1979; Strasser and Delson, 1987). These categories were accepted until early molecular studies suggested that the mangabeys may not constitute a monophyletic group (Cronin and Sarich, 1976; Disotell et al., 1992; Disotell, 1994; Harris and Disotell, 1998; Tosi et al., 1999; Disotell, 2000; Harris, 2000; Tosi et al., 2003). In particular, one subset of mangabeys (now called *Lophocebus*) was suggested to be more closely related to baboons (*Papio*), and the

other (*Cercocebus*) to drills and mandrills (*Mandrillus*). Dental and postcranial similarities among the *Cercocebus*/*Mandrillus* and *Lophocebus*/*Papio* pairs were uncovered shortly thereafter, further supporting the idea of paraphyly in this clade (Fleagle and McGraw, 1999; 2002; McGraw and Fleagle, 2006; Gilbert, 2007).

Many papionin species are characterized by extreme sexual dimorphism, and considerable size differences exist among sexes and taxa (Jolly, 1972; Groves, 1978; Fleagle and McGraw, 1999). As such, some aspects of their cranial morphology appear homoplastic, until the influence of sexual dimorphism and body size allometry are thoroughly accounted for (Singleton, 2002; Frost et al., 2003; Gilbert and Rossie, 2007; Gilbert, 2008; 2009; Gilbert et al., 2009; Gilbert, 2011). Superficially, the small-bodied mangabeys (*Cercocebus* and *Lophocebus*) tend to resemble each other in cranial shape, while the large-bodied genera (*Mandrillus*, *Papio*, and *Theropithecus*) tend to share a general cranial form (Thorington and Groves, 1970; Jolly, 1972; Szalay and Delson, 1979; Strasser and Delson, 1987). Several studies have attempted to correct for the homoplastic impact of allometry in this tribe with varying degrees of success (Frost et al., 2003; Gilbert and Rossie, 2007; Gilbert et al., 2009; Gilbert, 2011).

Gilbert and colleagues found that among the highly dimorphic papionin species, the morphology of male individuals tends to reflect phylogeny more accurately than that of females, and suggested that the sexes be analyzed separately in subsequent analyses (Gilbert and Rossie, 2007; Gilbert et al., 2009). Narrow allometric coding (Gilbert and Rossie, 2007) and general allometric coding (Gilbert et al., 2009) have both been applied to traditional craniometric characters in an attempt to control for the effects of allometry in cladistic analyses of papionin primates. These allometric corrections resulted in cladograms that were consistent with the molecular phylogeny at the genus level (Gilbert and Rossie, 2007; Gilbert et al., 2009). However, these studies did not attempt to resolve interspecific relationships below the genus level.

Applying allometric adjustment techniques to 3D cranial landmark data in geometric morphometric analyses proves more computationally and theoretically complicated (Frost et al., 2003; Gilbert, 2011). Recently, Gilbert attempted a size correction technique on the 3D basicranial morphology of African papionin primates (Gilbert, 2011). In this approach, all principal components (PCs) resulting from the geometric morphometric analysis were regressed against

centroid size, and any PCs that were found to be correlated with centroid size were then eliminated from the analysis. However, this approach did not yield trees that were congruent with the molecular phylogeny, and Gilbert explicitly indicated the need for an improved method of allometric correction in this group (Gilbert, 2011). As acknowledged by Gilbert, an additional potential limitation of this technique is that when size-correlated PCs are entirely excluded from an analysis, additional non size-correlated information about shape contained therein will necessarily be removed along with them. Gilbert calculated his thresholds for PC removal at r=0.576 for his combined sex analyses and r=0.881 for the individual sex analyses. Even using the higher critical r-value could result in up to 23% of the variance explained by an eliminated PC to be unrelated to size, i.e., pure shape information. Thus, omitting these size-correlated PCs has the potential to result in the exclusion of a substantial amount of shape information along with them. Another approach to allometric adjustments in papionin cranial datasets was described by Frost and colleagues (Frost et al., 2003). In this method, the geometric morphometric shape coordinates are regressed against centroid size, and the resulting residuals are extracted from the regression as new "size-corrected" variables (Frost et al., 2003). This technique effectively enables the factoring out much of the portion of the shape coordinates that is explained by size, resulting in allometrically corrected data. However, some shape information is also eliminated along with scaling information using this method (Jungers et al., 1995; Gilbert, 2011).

Given the foregoing discussion, this study has two interconnected aims. Firstly we explicitly test the genetic congruence of various cranial regions delineated according to the three main criteria described in detail above: Biomechanical strain, development and (presumed) heritability, and anatomical and functional complexity. Use of these criteria allowed us to further assess the validity of the theoretical assumptions underlying these criteria, and also provided an empirical basis for delineating cranial units that adhere to previously hypothesized predictions regarding phylogenetic utility. In order to do so we statistically compare the relative strength of correlation between a genetic distance matrix representing the genetic relationships among papionin taxa and various morphological distance matrices based on different aspects of cranial morphology. Throughout the paper we refer to these correlations as (relative) "genetic congruence". Our study was explicitly designed to facilitate direct comparison with the results of a previous study based on hominoid taxa (von Cramon-Taubadel and Smith, 2012) as the landmark

configurations employed in these studies are intentionally identical in configuration. This allowed us to assess whether the cranial regions found to yield the strongest correlations with genetic distances in hominoids are also those most genetically congruent in papionins. Any agreement or disagreement in terms of *patterns* of genetic correlations may provide important insight into the potential phylogenetic utility of various cranial regions across fossil primates irrespective of the theoretical criteria employed to delineate them. In addition, since papionin cranial morphology is heavily impacted by the effects of sexual dimorphism and allometry, our second aim was to assess the effects of both separation of the sex samples and a computational method of 3D allometric adjustment on the initial genetic congruence results obtained.

Materials and methods

Materials: Morphological

Landmark configurations

Three-dimensional coordinates of 147 craniomandibular landmarks (Table 1; Figs. 1-4) were digitized by one of us (HFS) using a Microscribe digitizer (Immersion Corp.). Landmark configurations were derived from previous studies on cranial morphology in human populations (Smith et al., 2007; Smith, 2009; von Cramon-Taubadel, 2009a; b; 2011b) and hominoid taxa (von Cramon-Taubadel and Smith, 2012). Following von Cramon-Taubadel and Smith (2012) and in order to maintain consistency with that study, for specimens with a prominent sagittal crest, it was necessary to transpose midline cranial vault landmarks (bregma and lambda) onto the anatomical midline. This transformation was achieved by orthogonally projecting these landmarks onto a midsagittal plane as defined by the midline landmarks glabella, prosthion, and basion (von Cramon-Taubadel and Smith, 2012).

The landmarks were divided into various subsets representing cranial regions as delineated in previous studies according to the three main criteria discussed earlier (Table 1, Figs. 1-4). Specifically, three developmentally-delineated regions were recognized: chondrocranium, neurocranium, and splanchnocranium (Fig. 1). Cranial regions representing areas presumed to be under high and low levels of masticatory strain were included: high strain= mandible, palatomaxilla and zygotemporal; low strain= basicranium, neurocranium, and upper face (Fig. 2). Finally, landmark subsets depicting the shape of eight individual cranial bones were delineated:

frontal, maxilla, occipital, parietal, sphenoid, temporal, and zygomatic bones (Fig. 3). It is important to note that the landmark configurations employed and the anatomical regions delineated are the same as those employed in a previous study of hominoid taxa (von Cramon-Taubadel and Smith, 2012).

Samples

A sample of 370 adult specimens comprising 14 cercopithecoid taxa were included in this study (Table 2). All specimens measured were non-pathological and fully adult as defined by a fused spheno-occipital synchondrosis. While there has been considerable discussion regarding the taxonomic assignment of the various *Papio* allotaxa (Jolly, 1993; Groves, 2001; Jolly, 2001), this level of systematic distinction is not crucial in the present study. Since taxa here are differentiated by genetic distances, whether the *Papio* allotaxa should be categorized as subspecies or separate species is immaterial to this study's research questions. Here we use the species-level designation.

Previous studies have revealed notable sex differences in the patterns of cranial morphology between male and female papionin primates (Fleagle and McGraw, 1999; 2002; Frost et al., 2003; McGraw and Fleagle, 2006; Gilbert, 2007; Gilbert et al., 2009). Thus, in order to determine whether the degree of correspondence between cranial matrices and the genetic distance matrix varied between the sexes, we also conducted separate sex analyses. Of the 14 species included in this study, four were found to have non-equivalent sex samples: *Macaca mulatta*, *Papio cynocephalus*, *P. hamadryas*, and *P. papio* (Table 2). The sex-bias in these samples precluded these taxa from being included in sex-separated analyses due to a lack of an acceptable sample size for one or both sexes. Therefore, a reduced 10 taxon dataset was utilized to conduct the separate sex analyses.

Thus, in sum, morphological distance matrices based on each of the 15 cranial regions were compared to the genetic distance matrix four times: once with the entire combined-sex 14 taxon sample, and again with only the 10 taxa with comparable sex compositions, the latter of which was run separately as combined-sex, male-only, and female-only iterations.

Materials: Genetic distances

Information concerning the consensus genetic relationships of the 14 papionin taxa was collected from previously published literature (Newman et al., 2004; McGoogan et al., 2007; Perelman et al., 2011; Springer et al., 2012). Most taxa considered in the present study were included in a recent high-resolution molecular phylogeny of living primates (Perelman et al., 2011). The Perelman et al. phylogeny is quite comprehensive, recent, and based on more than 30,000bp of nuclear DNA; thus, we used its topology and associated branch lengths as a baseline. However, three of the papionin taxa in the present study were not included in the phylogeny of Perelman and colleagues (2011). Thus, we drew upon the results of other phylogenetic studies (Newman et al., 2004; McGoogan et al., 2007) in order to resolve the complete set of genetic distances:

1. Perelman and colleagues (2011) included just three of the five allotaxa of *Papio* in their analysis. Thus, the genetic distances for *Papio cynocephalus* and *P. ursinus* were calculated from supplementary published sources. Newman et al. (2004) and McGoogan et al. (2007) both positioned *P. ursinus* as the outgroup to the rest of the *Papio* clade, and *P. cynocephalus* as the sister species to *P. anubis*. Thus, this topology was adopted here. Pairwise genetic distances between *Papio* taxa presented by Newman et al. were used to estimate distances between each species and the added nodes (15a, 17a), and scaled to the *Papio* branch lengths reported in Perelman et al. (2011). Due to the extremely close relationship between *P. anubis* and *P. cynocephalus*, indicated by a low genetic distance in Newman et al. (2004), these species were treated as equidistant from node 15a. Positioned as the sister to the rest of the *Papio* clade, the branch length for *P. ursinus* was estimated by employing its relative genetic distance from them (from Newman et al., 2004) and scaling it to the distances between the next two outside members of the clade, *P. papio* and *P. hamadryas*.

2. Perelman et al. (2011) included *Cercocebus agilis and C. torquatus,* but not *Cercocebus atys* in their analysis. Given that until recently, *C. atys* was subsumed within *C. torquatus* (removed by Groves 2001), and the fact that they were found to be sister taxa by Springer and colleagues (2012) and McGoogan and colleagues (2007), we inferred a close phylogenetic relationship between these taxa, and treated them as equidistant from their

common node (21a). This recent taxonomic designation means that little data exist on the relative distance of *C. atys* to other species. To estimate the distance between *C. atys* and *C. torquatus*, we scaled the branch length of *C. torquatus* from Perelman et al. (2011) by the phylogenetic diversity between *C. atys* and *C. torquatus* indicated by McGoogan et al. (2007).

Figure 5 illustrates the complete set of branch lengths derived from the literature. In order to transform the consensus phylogeny into a genetic distance matrix that could be statistically compared to the cranial data matrices, we generated a pairwise genetic distance matrix among all taxa (Cardini and Elton, 2008). Following the procedure described by Cardini and Elton (2008), we converted the above-described phylogenetic branch lengths into a genetic distance matrix for all papionin taxa (Figure 5). Using Phylip 3.695, Neighbor-joining trees (Saitou and Nei, 1987) were then generated from the distance matrix to confirm that the published topology was correctly recovered from the genetic distance matrix.

Analytical methods

Morphological distance matrices

In order to evaluate the relative strength of statistical correlations between each of the 15 cranial regions and the genetic distance matrix, various morphological distance matrices were generated according to the following procedure: First, landmark coordinates from each cranial subset were superimposed using Generalized Procrustes Analyses (Gower, 1975) in *MorphoJ 1.05f* (Klingenberg, 2011). Mahalanobis D^2 distances between each pair of taxa were then calculated from the resultant Procrustes variables, and entered into a matrix of pairwise morphological distances. This procedure was repeated separately for the subset of landmarks representing each cranial region (Table 1) for each dataset (the entire combined-sex 14 taxon dataset, reduced 10 taxon combined-sex dataset, 10 taxon male-only dataset, and 10 taxon female-only dataset).

Comparing morphometric and genetic distance matrices

Each morphological distance matrix was statistically compared to the genetic distance matrix using a Mantel test (Mantel, 1967) in PopTools, an add-on for Microsoft Excel. Matrix permutations (10,000 iterations) were then used to assess significance, with the alpha level set at $\alpha = 0.05$. Subsequently, Dow-Cheverud tests (Dow and Cheverud, 1985) were conducted to

determine whether any of the morphological distance matrices were more strongly correlated with the genetic distance matrix than were others. Dow-Cheverud tests were executed in *R* using a script written by the lab of Charles Roseman. The strength of this test was assessed using the variable p1Z and its corresponding *p*-value (with critical alpha set at <0.05). The p1Z value is the statistical value of the difference in correlation of two matrices with a third; in this case, the difference in correlations between two morphological matrices (representing two cranial units) and the genetic distance matrix. All Mantel and Dow-Cheverud tests were performed separately for each of the 15 cranial regions in each of the four data sets. For each set of analyses, a sequential Bonferroni correction was also applied to correct for multiple comparisons (Holm, 1979).

Allometric scaling adjustments

Numerous studies have described allometric patterning in papionin cranial morphology (Fleagle and McGraw, 1999; 2002; Frost et al., 2003; Gilbert, 2007; 2008; 2009; Gilbert et al., 2009), the result of which can obscure the phylogenetic signal in the skull of this tribe (Gilbert and Rossie, 2007; Gilbert et al., 2009; Gilbert, 2011). While considerable discussion has centered around how to adjust for unwanted allometric cranial characteristics to reveal underlying phylogenetic and taxonomic information (Frost et al., 2003; Gilbert and Rossie, 2007; Gilbert et al., 2009; Gilbert, 2011), no consensus exists as to the most effective method of allometric adjustment, particularly in the case of 3D geometric morphometric data.

In the present study, we wanted to assess whether an allometric adjustment would increase the correlations between cranial morphological matrices and genetic distance matrices in papionin primates. To achieve this, we chose the allometric adjustment method described by Frost and colleagues (2003). After conducting the Generalized Procrustes Analysis (GPA) for each cranial region, we then performed a regression analysis in *MorphoJ* in which the Procrustes residuals were regressed against the natural log of centroid size. The residuals from this analysis were then extracted and used as allometric size-corrected variables, which were subsequently used to calculate a new set of allometric size-corrected Mahalanobis D^2 distances among taxa. This procedure was repeated independently for each cranial region/data set combination, resulting in a total of 60 additional allometric size-adjusted matrices (15 cranial regions in four data sets). Each allometrically adjusted matrix was then statistically compared to the genetic matrix using a Mantel test and Dow-Cheverud tests (as described above) to determine which allometric size-adjusted

cranial regions were significantly correlated with genetic distances for each data set. Finally, the results of the allometric size-corrected analyses were compared to those of the non size-corrected results using Dow-Cheverud tests to determine whether conducting an allometric adjustment in this manner resulted in significantly greater correlation coefficients between cranial morphological matrices and the genetic distance matrix.

Results

Mantel tests

Table 3 lists the results of all Mantel tests conducted between the genetic matrix and the various morphological matrices. Cranial regions are listed by category: functional-developmental units, individual cranial bones, and regions defined by degree of presumed masticatory stress. Each cranial region is labeled according to predictions made about its genetic congruence based on previous results from a similar analysis of hominoid taxa (von Cramon-Taubadel and Smith, 2012), with 1 being statistically more genetically congruent than 2, which is statistically more congruent than 3.

Complete 14 taxon combined-sex data set

The vast majority of the Mantel tests between the cranial region matrices and genetic matrix for the complete 14 taxon data set yielded significant correlations (Table 3). Only the parietal and sphenoid matrices were not found to be significantly correlated with genetic distances. The chondrocranium and zygotemporal matrices resulted in the highest correlations with the genetic matrix, both of which were absolutely higher than that based on the shape of the entire cranium (Table 3). All of these correlations remained significant even following application of sequential Bonferroni adjustment.

Reduced 10 taxon data set

The correlations obtained in the Mantel tests for the reduced 10 taxon data set were, on average, lower than those of the 14 taxon data set (Table 3). Of the functional-developmental cranial regions, the chondrocranium and neurocranium matrices were significantly correlated with the genetic distance matrix (Table 3). Of the masticatory strain regions, the correlations for the

zygotemporal and mandibular matrices were also significant, as were the correlations based on the temporal, zygomatic, and occipital bone matrices (Table 3). However, only the chondrocranium matrix remained significantly correlated with the genetic distance matrix following sequential Bonferroni adjustment.

The separated sex analyses yielded fewer significant correlations than the combined-sex analysis for the reduced 10 taxon data set (Table 3). For the male-only comparisons, only the chondrocranium and the temporal bone matrices were significantly correlated with the genetic matrix following sequential Bonferroni adjustment (Table 3). In the female-only analysis, only the zygotemporal and occipital bone matrices were found to be significantly correlated with the genetic matrix. However, none of the female-only Mantel tests were significant following application of sequential Bonferroni adjustment (Table 3).

Dow-Cheverud tests

Complete 14 taxon combined-sex data set

A few cranial matrices in the complete 14 taxon data set were found to differ significantly in their correlation with the genetic distance matrix (Table 4). Most notably, the chondrocranium matrix was significantly more genetically congruent than the other two developmental matrices and the upper face matrix. The sphenoid and parietal bone matrices were the least correlated with the genetic matrix, significantly less than almost all other individual bone matrices (Table 4). There were very few differences in terms of genetic congruence among the remaining cranial matrices (Table 4). Moreover, when the stricter alpha-levels associated with the sequential Bonferroni correction are applied, the only significant Dow-Cheverud test result was that the sphenoid matrix was less genetically congruent than the matrices based on the entire cranium, maxilla, temporal, and occipital bones.

Reduced 10 taxon data set

The reduced 10 taxon data set revealed fewer significant differences among the cranial regions in terms of their genetic congruence (Table 5). For the combined-sex analysis, the chondrocranium matrix was significantly more genetically congruent than the splanchnocranium matrix and the matrix based on the entire cranium (Table 5). The temporal and zygomatic bone matrices were

both significantly more genetically congruent than the matrices based on the maxilla, sphenoid and the entire cranium (Table 5). Of the regions associated with masticatory strain, the chondrocranium matrix was significantly more genetically congruent than the mandible, palatomaxilla, upper face, and entire cranium matrices, and the zygotemporal region was significantly more genetically congruent than the upper face and entire cranium. However, as in the case of the complete 14 taxon dataset, application of a sequential Bonferroni correction yielded few significant correlations. In this case the temporal and occipital bone matrices were significantly more genetically congruent than the sphenoid bone matrix (Table 5).

The Dow-Cheverud tests for the separated sex analyses yielded a smaller number of significant differences in correlations between cranial matrices and the genetic matrix. For the male-only data set (Suppl. Table 1), the chondrocranium matrix was significantly more genetically congruent than the other developmental region matrices even following sequential Bonferroni correction. Also, the temporal bone matrix was more strongly correlated with the genetic matrix than all bone matrices, except for the zygomatic matrix, although these results were not significant following Bonferroni correction. The zygotemporal matrix was also more genetically congruent than the neurocranium matrix and the entire cranium matrix, but not once sequential Bonferroni correction was applied. In the female-only analysis (Suppl. Table 2), the only significant Dow-Cheverud tests were that the zygomatic and the entire cranium matrix were more genetically congruent than the maxilla matrix, and that the zygotemporal matrix was more genetically congruent than the chondrocranium matrix. However, these Dow-Cheverud comparisons were not found to be significant following sequential Bonferroni adjustment of alpha-levels.

Allometric size-adjusted analyses

Performing allometric adjustments on the morphological data resulted in some increased correlations between the cranial distance matrices and the genetic distance matrix (Table 6). This improved correlation was most pronounced in the male-only data set, and least prevalent in the combined-sex 14 taxon data set, where the majority of the morphological matrices were already significantly correlated with the genetic distance matrix prior to the allometric size-correction.

<u>Complete 14 taxon combined-sex data set</u>

The allometrically adjusted matrices for the 14 taxon data set were on average found to be quite similar to those of the uncorrected matrices in terms of patterns of genetic congruence (Table 6). The same thirteen morphological matrices were found to be significantly correlated with the genetic matrix, such that even after the allometric size-correction, only the parietal and sphenoid bones remained non-significantly correlated with the genetic matrix. On average, correlation coefficients across the cranial regions actually decreased following the allometric correction. Dow-Cheverud tests examining whether the allometrically-adjustment resulted in significant improvements over the non-adjusted data did not indicate any significant improvements by adding the size-correction for the 14 taxon data set (Suppl. Table 3).

Dow-Cheverud tests revealed that the allometric size-adjustment resulted in the chondrocranium matrix being significantly more congruent with the genetic distance matrix than the entire cranium matrix, although this finding did not hold up to a sequential Bonferroni adjustment. As before, the parietal and sphenoid bone matrices were significantly less genetically congruent than several other individual bone matrices (Table 7). Following sequential Bonferroni adjustment, the only significantly different result that remained were that the sphenoid bone matrix was significantly less genetically congruent than the occipital bone matrix.

Reduced 10 taxon data set

The 10 taxon dataset demonstrated a greater number of increased *r*-values and significant correlations after the allometric correction (Table 6). For the combined-sex data set, the splanchnocranium, palatomaxilla, and upper face matrices were significantly correlated with the genetic matrix following the allometric size-adjustment. The chondrocranium, neurocranium, zygotemporal, mandible, zygomatic, temporal, and occipital matrices remained significantly correlated with the genetic matrix, and all yielded an absolutely higher *r*-value than in the uncorrected comparison (Table 6). In the Dow-Cheverud tests comparing correlations before and after the allometric size-correction, only the splanchnocranium and sphenoid matrices demonstrated significantly increased correlations with the genetic distance matrix following the allometric adjustment (Suppl. Table 3).

Dow-Cheverud tests also indicated that the increased genetic congruence of the splanchnocranium and neurocranium matrices following the allometric size correction meant that

these matrices are significantly more strongly correlated with the genetic matrix than was the entire cranium (Table 8). None of the masticatory strain region matrices differed significantly from each other in terms of genetic congruence. However, the chondrocranium, palatomaxilla, and zygotemporal matrices were still significantly more correlated with the genetic matrix than the entire cranium matrix.

In the allometrically adjusted male-only analyses, the correlation coefficients increased more dramatically than in any other allometric size-corrected data set (Table 6). Of the developmental regions, both the splanchnocranium and neurocranium matrices were significantly correlated with the genetic matrix after the size-correction, while the chondrocranium matrix remained significantly correlated (Table 6). Four additional masticatory strain regions were found to yield genetically congruent matrices following the allometric size-correction: the mandible, palatomaxilla, upper face, and zygotemporal matrices (Table 6). The maxilla bone matrix was also significantly correlated with the genetic matrix following the allometric size-adjustment procedure (Table 6). Dow-Cheverud tests comparing cranial-genetic matrix congruence before and after the allometric size-correction indicated that the palatomaxilla matrix was significantly more genetically congruent following this adjustment (Suppl. Table 3). Dow-Cheverud tests in the male-only size-adjusted analysis revealed no significant differences in terms of genetic congruence among any of the developmental or masticatory strain matrices (Table 9). Among the individual cranial bone matrices, the temporal bone matrix was significantly more genetically congruent than the frontal, parietal, sphenoid, and entire cranium matrices. The entire cranium and the zygomatic bone matrices were both significantly more highly correlated with the genetic matrix than the sphenoid matrix. However, no individual comparisons remained significant following sequential Bonferroni adjustment (Table 9).

Following the allometric size-adjustment, the female-only data set continued to display the fewest number of significant correlations between morphological and genetic distance matrices all the data sets (Tables 6). The entire cranium, splanchnocranium, upper face, zygomatic, and temporal bone matrices were significantly correlated with the genetic matrix after the size-correction. The occipital bone matrix continued to reflect genetic distances as in the uncorrected analysis; however, the previously genetically congruent zygotemporal matrix was not significantly

correlated with the genetic matrix (Table 6). Dow-Cheverud tests revealed that the correlation between the splanchnocranium matrix and genetic matrix was significantly increased by the size-adjustment (Suppl. Table 3).

The Dow-Cheverud tests for the female-only analysis revealed significant differences in genetic congruence among a few cranial matrices after the allometric size-adjustment (Table 10). Of the developmental regions, the splanchnocranium matrix was significantly more strongly correlated with the genetic matrix than either the chondrocranium or neurocranium matrices (Table 10). The entire cranium matrix was significantly more genetically congruent than the neurocranium, mandible, maxilla, sphenoid, and zygotemporal matrices. The occipital and temporal bone matrices were found to be significantly more genetically congruent than the maxilla matrix. The decrease in the Mantel test correlation between the zygotemporal matrix and the genetic matrix resulted in this matrix no longer being significantly more genetically congruent than the neurocranium, mandible and chondrocranium matrices (Table 10), when compared to the non-adjusted analysis. However, as was the case in the male-only analysis, no individual Dow-Cheverud comparison was significant following sequential Bonferroni correction.

The final set of Dow-Cheverud analyses performed were designed to assess whether the sex-specific plus allometrically size-adjusted datasets yielded significantly more genetically congruent cranial data matrices than combining the sexes together (Table 11). The results show that in the case of males, the occipital bone matrix was more genetically congruent in the combined-sex dataset compared with analyzing males separately (Table 11). In the case of females, the entire cranium and the sphenoid matrices were more genetically congruent in the female-only dataset than in the combined sex sample. Several regions, including the chondrocranium, neurocranium, palatomaxilla, zygotemporal, temporal, maxilla and the mandible matrices were more highly correlated with the genetic matrix in the combined-sex sample than in females alone.

Discussion

The results presented here illustrate the lack of any single criterion that can be used to delineate ubiquitously genetically congruent cranial modules among papionin taxa. Additionally, the relative

correlations of cranial region matrices and the genetic matrix differed somewhat between the sexes, while applying an allometric correction to the sex-separated data resulted in some additional cranial matrices being significantly correlated with genetic distances in each sex. In both the full 14-taxon and reduced taxon combined-sex data sets, both the high masticatory strain regions of the mandible and zygotemporal, as well as the low strain chondrocranium and neurocranium yielded matrices that were found to be correlated with the genetic matrix (Table 3). However, none of the various strain region matrices was found to be significantly more genetically congruent than any other, lending no support to the homoiology hypothesis (Tables 4, 5, 7 and 8). Of the developmental regions, the chondrocranium matrix was consistently more correlated with genetic distance than the neurocranium and splanchnocranium matrices, and sometimes significantly more genetically congruent than the entire cranium matrix, when sexes were combined (Tables 4, 5, 7). Taken at face value, this might suggest some support for the "basicranial hypothesis" in mixed-sex (or unknown sex) analyses. However, it should be noted that in cases where the stricter sequential Bonferroni adjustment was applied there was no statistical difference in the genetic congruence of the chondrocranium matrix and any other cranial region matrices, suggesting that support for this hypothesis is weak at best. The anatomical/functional complexity hypothesis received mixed support in the combined-sex data. The matrices based on the anatomically complex temporal bone were correlated with the genetic matrix in all combined-sex comparisons; however, they were often no more strongly correlated than other matrices based on relatively simple bones such as the occipital or the frontal. Additionally, another anatomically complex bone – the sphenoid – consistently yielded matrices with the lowest correlations with the genetic matrix (Tables 3, 6) across all analyses. Taken together, the results of the papionin analyses presented here do not support the internal predictions of the three major hypotheses regarding relative phylogenetic efficacy of cranial regions outlined in the Introduction.

Sexual dimorphism

In the case of the allometrically corrected datasets, partitioning the data by sex provided no improvement in terms of the relative strength of correlation between morphological matrices and the genetic matrix (Table 11). Basing morphological matrices on the male-only dataset resulted in very similar genetic correlations to those obtained using the combined-sex dataset. Conversely, the

matrices based on the chondrocranium, neurocranium, maxilla, temporal bone, palatomaxilla, zygotemporal and mandible were all significantly more strongly correlated with the genetic matrix using the combined-sex database than they were using the female-only dataset. This result confirms the general observation (Tables 3 and 6) that the female-only dataset consistently yielded fewer morphological matrices that were significantly correlated with the genetic matrix, compared with the male-only dataset and the combined-sex dataset. The lower number of significant correlations between cranial matrices and genetic distance in female papionins is consistent with previous studies that have revealed a dearth of phylogenetic information in the cranial morphology of females from these species (Gilbert et al., 2009; Gilbert, 2011). In fact, it has previously been advocated that female papionin primates be analyzed separately from males, and that their cranial form may be less phylogenetically informative than that of their male counterparts (Gilbert et al., 2009). It has been suggested that male papionin cranial anatomy may reflect actual synapomorphic traits, (Gilbert et al., 2009), while female papionins tend to share a more generalized and undifferentiated cranial form. While our results show that females differ from males in terms of the extent to which morphological matrices are correlated with genetic distances, they do not provide compelling evidence in favor of separating the sexes in future studies, as the male-only dataset did not yield any morphological matrices that were significantly more genetically congruent than analogous matrices based on the morphology of both sexes.

Consequences of allometric size adjustment

The present study employed a method of allometric correction that involved regressing Procrustes residuals against centroid size and then taking the residuals of this analysis as new size-adjusted variables, following Frost and colleagues (2003). In the case of the larger 14-taxon dataset, the allometric adjustment had no effect on the pattern of Mantel test correlations between morphological matrices and the genetic distance matrix. In the case of the reduced 10-taxon dataset a few additional cranial matrices that centered on facial morphology were significantly correlated with the genetic distance matrix following the allometric adjustment. Following allometric adjustment, 11 out of the 15 cranial region matrices tested were significantly correlated with the genetic matrix in males. However, the correlation coefficients of the Mantel tests performed here were not high ($r = 0.235$-0.453), meaning that the majority of the morphological

variation is inconsistent with the genetic distances, and therefore unexplained by genetic relatedness. These allometrically-corrected values are much lower than many obtained for similar comparisons in hominoids (von Cramon-Taubadel and Smith, 2012) and humans (von Cramon-Taubadel, 2009a; b; 2011b). Additionally, Dow-Cheverud tests comparing the genetic congruence of morphological matrices before and after allometric adjustment indicated that only a few of these changes were statistically significant. Interestingly, all of the significantly different changes are centered on facial morphology (10-taxon dataset: splanchnocranium; male-only: palatomaxilla, female-only: splanchnocranium), which is consistent with existing knowledge regarding the homoplastic facial elongation that characterize the large-bodied papionin genera (*Papio* and *Mandrillus*) versus the short-faced *Lophocebus* and *Cercocebus* genera. Therefore, performing allometric corrections may allow for a better correlation between papionin facial morphological matrices and known genetic distances, yet may have little to no effect on the genetic congruence of non-facial morphological regions.

Comparison with patterns found in hominoid taxa

Despite the fact that the empirical findings do not lend support to the internal predictions of the outlined hypotheses, we can compare the patterns of genetic congruence found here directly with those found in an analogous study based on hominoid taxa (von Cramon-Taubadel and Smith, 2012). The pattern of results obtained here differs greatly from that found in hominoids, despite utilizing the same landmark definitions to categorize each cranial region and the same analytical approach. In the case of the hominoids (von Cramon-Taubadel and Smith, 2012), the morphological distance matrices for all cranial regions were significantly correlated with the genetic distance matrix. In the latter study sequential Bonferroni correction was not employed, but it is worth noting that the highest p-value obtained was 0.002 so these correlations would have been considered significant even utilizing a strict Bonferroni correction ($\alpha = 0.0027$). Moreover, as noted above, the *r*-values were overall much higher in the hominoid study (i.e. ranging from $r = 0.35$ to 0.86), suggesting that, in general, the correlation between cranial distances and genetic distances is stronger in hominoids than in papionins. This result is also consistent with the findings of Roseman et al. (2010) who found no differences in the heritability or conditional evolvability of different cranial regions in a captive population of baboons. This suggests that there may be few

differences across the papionin cranium in terms of phylogenetic efficacy, with all regions being equally useful (or problematic) when attempting to reconstruct phylogenetic relatedness across taxa. This finding is consistent with previous phylogenetic studies implementing various alternate methods of size correction (Gilbert and Rossie, 2007; Gilbert et al., 2009; Gilbert, 2011), and suggests that cranial morphology in papionins, while containing some degree of information about genetic relatedness, is simply not as informative as in other primate taxonomic groups.

The relative pattern of correspondence between morphological and genetic distance for different cranial regions (i.e., within each of the delineated cranial categories: biomechanical strain, development and heritability, and complexity) also differed between the hominoid and the papionin datasets. Taking each of the criteria in turn: In the case of the biomechanical strain criterion, the hominoid pattern revealed that the morphological/genetic correlations were significantly higher for the upper face and the zygotemporal regions than for the neurocranium and chondrocranium, which in turn were higher than the mandible and palatomaxilla (von Cramon-Taubadel and Smith, 2012). These findings did not match the specific predictions of the homoiology hypothesis, which would predict significantly higher correlations for the non-masticatory regions of the upper face, neurocranium, and cranial base relative to the masticatory regions of the palatomaxilla, mandible, and zygotemporal. In the case of the present study, the internal predictions of the homoiology hypothesis were also not met. Moreover, the relative pattern of correspondence between morphological and genetic distances found in papionins differed from that in the hominoids with the chondrocranium matrix generally having significantly higher correlations than the upper face and many of the masticatory regions.

In the case of the second criterion of development and heritability, the hominoid dataset did not follow the predictions of the "basicranial hypothesis" in that morphological/genetic congruence was consistently greater in the splanchnocranium than it was in either the chondrocranium (basicranium) or the neurocranium (von Cramon-Taubadel and Smith, 2012). In the case of the papionin datasets analyzed here, the chondrocranium matrix was consistently one of the most genetically congruent of all the cranial matrices tested and was generally found to have statistically higher correlations with the genetic matrix than the face or the cranial vault matrices across all datasets tested. This is obviously in contrast with the hominoid findings but does mirror

more closely the findings of Cardini and Elton (2008) who found the shape of the guenon chondrocranium yielded distance matrices more strongly correlated with the genetic distance matrix than other major cranial regions. This hints at the fact that the chondrocranium may offer a more accurate reflection of the underlying phylogenetic history of the cercopithecoids compared with other regions such as the splanchnocranium and neurocranium, while in the case of hominoids the facial skeleton may reflect among-taxon relationships more accurately.

Finally, in the case of individual bones, the hominoid study (von Cramon-Taubadel and Smith, 2012) found that morphological distances based on the zygomatic best mirrored genetic distance, followed by the neurocranial bones (frontal, parietal, and sphenoid), with the temporal and occipital matrices showing the lowest correspondence between morphological and genetic distances. As outlined previously this also does not match the internal predictions of the anatomical/functional complexity hypothesis. In the case of the papionins, the pattern was somewhat different with the temporal and occipital bones generally showing significantly higher correlations between morphological and genetic distance than other bones, and the parietal and sphenoid matrices generally showing the lowest correlations. One point of overlap between the hominoid and papionin results is the overall pattern of generally high genetic congruence for the zygomatic bone matrix, across all datasets and both sexes. Nevertheless, the papionin results did not match the internal predictions of the anatomical/functional complexity hypothesis nor did they match the pattern of genetic congruence found in hominoids.

It is also worth noting that in the case of the hominoids, despite substantial sexual dimorphism among some taxa, the relative correlations between the genetic distance matrix and cranial matrices were very similar across the sex-specific analyses and the combined-sex dataset (von Cramon-Taubadel and Smith, 2012). However, in the case of the papionins, the sex-specific analyses yielded somewhat different patterns both between the sexes and relative to the combined-sex dataset. In particular, morphological matrices based on the female papionin dataset yielded significantly weaker correlations with the genetic matrix than the male-only or the combined datasets. This is consistent with the relationship between sexual dimorphism and phylogenetic reconstruction that has been noted so often for papionin taxa (Singleton, 2002; Frost et al., 2003; Gilbert and Rossie, 2007; Gilbert, 2008; 2009; Gilbert et al., 2009; Gilbert, 2011). One possible

explanation for this may be the effect of genetic variance and sex interaction effects concentrated in the neurocranium and facial skeleton of papionins (Willmore et al., 2009), despite an approximately equivalent level of heritability across traits from all cranial regions. If it is the case that sex-specific quantitative genetic patterns in papionins are concentrated in the anatomical junction between the face and the neurocranium, this may also help explain why morphological distances for the chondrocranium were found to be the most reliably correlated with genetic distance in the combined-sex analyses.

In sum, there is no consistent criterion that appears to explain relative genetic congruence of cranial regions across both hominoids and papionins. In hominoids, facial morphology appears to be the most reliable correlate of genetic distances in both males and females (von Cramon-Taubadel and Smith, 2012). Finally, as shown here and corroborated by previous studies (Gilbert, 2011), basicranial morphological distances are correlated with genetic distances among male papionin primates. The fact that the same pattern holds true for guenons (Cardini and Elton, 2008) suggests that this may be a cercopithecoid-specific pattern rather than something unique to the papionins. In the case of the papionins, an allometric size-adjustment of some form is necessary when analyzing the facial skeleton, and there is still debate about the most effective way to implement one (Frost et al., 2003; Gilbert and Rossie, 2007; Gilbert et al., 2009; Gilbert, 2011).

Conclusions and future considerations

Taken together these findings suggest that a single, broad model explaining or predicting cranial phylogenetic utility across all catarrhine primates is simply not tenable. The results obtained here and in the case of our previous work on hominoids (von Cramon-Taubadel and Smith, 2012) would suggest that future studies attempting to reconstruct phylogeny or taxonomy from fossil cranial specimens are probably best served by using all available cranial information. Nevertheless, given the fragmentary nature of the fossil record, it should be recognized that not all cranial fragments are likely to be equally useful in terms of reconstructing phylogenetic history across all taxonomic groups. In the case of the papionins, employing some type of allometric size-adjustment for facial morphology is likely to be important to counteract the homoplastic evolution characterizing the *Cercocebus*/*Mandrillus* and the *Lophocebus*/*Papio* clades, although it is equally important to recognize the limitations of such correction methods both in terms of the potential for

discarding relevant size and shape information and uncertainty regarding the applicability of such methods to fossil taxa.

This study also further highlights the importance of applying taxonomically- and phylogenetically-appropriate evolutionary models when assessing the congruence between morphological and genetic among-taxon distances. In previous studies comparing cranial distance matrices against neutral genetic distance matrices (e.g., Roseman, 2004; Harvati and Weaver, 2006; Smith, 2009; von Cramon-Taubadel, 2009a, 2009b, 2011a, 2011b), the theoretical assumption being made is that morphological matrices that correlate more strongly with neutral genetic distances are doing so because the history of among-population divergence for that particular morphology is largely due to neutral stochastic processes such as genetic drift. Deviations from proportionality between morphological and genetic among-population distance matrices could be indicative of diversifying selection or other non-neutral factors in certain aspects of morphology in particular populations (Roseman, 2004). When this theoretical assumption is scaled up to supra-specific taxonomic levels, the expectation is that among-taxon morphological distances scale linearly with time, in the same way that among-taxon genetic distances are expected to when genetic phylogenies are reconstructed using non-coding genomic data (e.g., Perelman et al., 2011). Thus, under a neutral model of morphological divergence among taxa (Lynch, 1990), we might expect morphological distance to reflect genetic distances. However, in reality, the speciation/divergence events that occurred to create the taxonomic diversity of the extant primates will include a combination of neutral divergence and specific instances of diversifying natural selection (Marroig and Cheverud, 2004), some of which may have led to specific convergences (i.e., homoplasy) in the cranial morphology of particular lineages. Given this, there are unlikely to be any set "rules" governing which aspects of an organism's phenotype or which parts of a phylogenetic tree are more/less likely to reflect such homoplastic events. This is made clear when we consider that the pattern of results obtained for hominoid and for papionin taxa are entirely different in terms of the overall strength of congruence between genetic and morphological distance matrices, and that none of the empirical expectations of previously proposed hypotheses concerning homoplasy are upheld in these two catarrhine groups. These observations are likely to reflect the different evolutionary histories of morphological diversification in these two primate groups.

Where do we go from here? We need to develop a better understanding of how evolutionary forces such as genetic drift and natural selection contribute to speciation and taxonomic diversification, which in turn will provide insight into specific instances of phenotypic convergence across particular lineages. One misconception that lies at the heart of hypotheses regarding homoplasy is the assumption that morphological traits with higher heritability will be more "reliable" for reconstructing phylogeny due to the notion that such traits reflect underlying genetic information more accurately and are less influenced by the environment (i.e., will be less plastic). These misconceptions have been discussed at length previously (von Cramon-Taubadel, 2009b; Roseman et al., 2010) and the problem lies with a general misunderstanding of how additive genetic variance and environmental variance combine and interact to produce observable phenotypic variance. Moreover, while the response to selection of a single quantitative trait is directly related to its heritability (i.e., the proportion of phenotypic variance due to underlying additive genetic variance), the relationship between heritability and the propensity for a multivariate morphological structure to respond to selection is more complex. Due to the effects of genetic and developmental pleiotropy (Cheverud, 1982; 1988; 1996) phenotypic traits are intercorrelated in systematic ways (i.e., integrated *sensu* Olson and Miller) (Olson and Miller, 1958) such that any potential response to selection of one trait is constrained by correlations with other traits under the influence of stabilizing selection (Lande, 1976; Lande and Arnold, 1983). Combining the factors of heritability, pleiotropy, and integrations leads to the concept of conditional evolvability or the ability of a phenotype to respond to directional selection, given constraints imposed by stabilizing selection on integrated traits (Hansen and Houle, 2008). Consideration of how conditional evolvability might operate highlights the main conceptual problem with the functional/anatomical complexity hypothesis: if these functional complexes are highly integrated, then a single selection event on that complex could, theoretically, result in homoplastic changes due to the high levels of covariance between different traits within the complex. Functional complexes (such as the basicranium) may also be relatively tightly integrated and subject to strong stabilizing selection, such that they vary less than one would expect under a neutral model of diversification. Therefore, both the ability to respond to directional selection (evolvability) and protection from diversifying selection (as mediated by stabilizing selection on integrated structures) could result in morphological lineages that have not diversified under neutral

rates of evolution, and therefore, would generate morphological distance matrices that do not correlate well with neutral genetic distance matrices.

Another issue to consider is the time-depth and the phylogenetic structure of the primate group under consideration. If one is comparing deep-time divergences among particular genera then an accurate reconstruction of this phylogeny may be best served by a dataset of relatively slowly evolving phenotypic traits, with a lower likelihood of multiple new (and potentially homoplastic) phenotypic variants having evolved across lineages. If, on the other hand, the aim is to reconstruct the phylogeny of a sample of relatively recently diverged and closely related taxa, then more rapidly evolving phenotypic traits that can capture these more recent radiations are more likely to be useful.

Recent studies have suggested that hominoids and cercopithecoids might differ in terms of their overall levels of skeletal integration and evolvability. Specifically, Young and colleagues (Young et al., 2010) demonstrated patterns of reduced integration among serially homologous limb segments in hominoids, while Old and New World monkey limb integration was much stronger, consistent with a bias in monkeys towards size-scaled variants of a basic quadrupedal *bauplan*. In contrast, the hominoid postcranial *bauplan* is highly variable reflecting a range of diverse locomotory repertoires include terrestrial knucklewalking, brachiation, and bipedalism. A similar pattern of strong integration in the cercopithecoid autopod, compared with a less integrated pattern in hominoids has also been shown (Rolian, 2009). Therefore, many of the differences we observe here in terms of patterns of morphological-genetic matrix correlations could be due to overall different levels of phenotypic integration and consequent evolvability in cercopithecoids and hominoids, although this remains to be explicitly tested for the cranium.

In sum, the results presented here add further empirical weight to the conclusion that there are no generalized predictors in terms of which aspects of primate cranial form are more/less likely to exhibit homoplasy. Further consideration of how morphological diversity patterns evolve at higher taxonomic levels within an evolutionary quantitative genetic framework makes clear why this is the case. Some primate taxa, as measured using skeletal phenotypic data, will have diversified under largely stochastic conditions, while other lineages will have diversified under the action of directional natural selection (Marroig and Cheverud, 2004). Some aspects of skeletal

anatomy (modules) have a higher propensity to evolve (both under neutral and selective conditions) while strong integration mediated by pleiotropy and stabilizing selection will limit the evolvability of other aspects of the phenotype. Disentangling and better understanding these processes and their consequences on the evolutionary history of the primates will serve to provide an enhanced inference model for understanding the evolutionary morphology of the fossil hominins.

Acknowledgements

We are grateful to Sarah Elton, Rebecca Ackermann, and three anonymous reviewers for their thoughtful comments on an earlier version of this manuscript. This research was funded by a grant from the Leakey Foundation. We thank Mark Grabowski for sharing his Dow-Cheverud *R* script with us, and Judy Chupasko (MCZ), Linda Gordon (NMNH), William Stanley (FMNH), and Eileen Westwig (AMNH), for permission to study collections in their care. All figures were generated by Brent Adrian.

References

Bjarnason, A., Chamberlain, A.T., Lockwood, C.A., 2011. A methodological investigation of hominoid craniodental morphology and phylogenetics. Journal of Human Evolution 60, 47-57.

Bouvier, M., 1986. A biomechanical analysis of mandibular scaling in Old World monkeys. American Journal of Physical Anthropology 69, 473-482.

Cardini, A., Elton, S., 2008. Does the skull carry a phylogenetic signal? Evolution and modularity in the guenons. Biological Journal of the Linnean Society 93, 813-834.

Cheverud, J.M., 1982. Phenotypic, genetic, and environmental morphological integration in the cranium. Evolution 36, 499-516.

Cheverud, J.M., 1988. A comparison of genetic and phenotypic correlations. Evolution 42, 958-968.

Cheverud, J.M., 1996. Developmental integration and the evolution of pleiotropy. American Zoologist 36, 44-50.

Collard, M., O'Higgins, P., 2001. Ontogeny and homoplasy in the papionin monkey face. Evolution and Development 3, 322-331.

Collard, M., Wood, B.A., 2007. Hominin homoiology: an assessment of the impact of phenotypic plasticity on phylogenetic analyses of humans and their fossil relatives. Journal of Human Evolution 52, 573-584.

Cronin, J.E., Sarich, V.M., 1976. Molecular evidence for dual origin of mangabeys among Old World monkeys. Nature 260, 700-702.

Disotell, T.R., 1994. Generic level relationships of the Papionini (Cercopithecoidea). American Journal of Physical Anthropology 94, 47-57.

Disotell, T.R., 2000. Molecular systematics of the Cercopithecidae. In: Whitehead, P.F., Jolly, C.J., editors. Old World Monkeys. Cambridge: Cambridge University Press. pp. 29-56.

Disotell, T.R., Honeycutt, R.L., Ruvole, M., 1992. Mitochondrial DNA phylogeny of the Old-World monkey tribe Papionini. Molecular Biology and Evolution 9, 1-13.
Dow, M.M., Cheverud, J.M., 1985. Comparison of distance matrices in studies of population structure and genetic microdifferentiation: quadratic assignment. American Journal of Physical Anthropology 68, 367-373.
Fleagle, J.G., McGraw, W.S., 1999. Skeletal and dental morphology supports diphyletic origin of baboons and mandrills. Proceedings of the National Academy of Sciences 96, 1157-1161.
Fleagle, J.G., McGraw, W.S., 2002. Skeletal and dental morphology of African papionins: unmasking a cryptic clade. Journal of Human Evolution 42, 267-292.
Frost, S.R., Marcus, L.F., Bookstein, F.L., Reddy, D.P., Delson, E., 2003. Cranial allometry, phylogeography, and systematics of large-bodied papionins (Primates: Cercopithecinae) inferred from geometric morphometric analysis of landmark data. The Anatomical Record Part A 275A, 1048-1072.
Gilbert, C.C., 2007. Craniomandibular morphology supporting the diphyletic origin of mangabeys and a new genus of the *Cercocebus*/*Mandrillus* clade, *Procercocebus*. Journal of Human Evolution 53, 69-102.
Gilbert, C.C., 2008. African papionin phylogenetic history and Plio-Pleistocene biogeography. Stony Brook, NY: Stony Brook University.
Gilbert, C.C., 2009. Phylogenetic history of the African papionins: a cladistic analysis of extant and fossil taxa using craniodental data. American Journal of Physical Anthropology 48S, 198.
Gilbert, C.C., 2011. Phylogenetic analysis of the African papionin basicranium using 3-D geometric morphometrics: The need for improved methods to account for allometric effects. American Journal of Physical Anthropology 144, 60-71.
Gilbert, C.C., Frost, S.R., Strait, D.S., 2009. Allometry, sexual dimorphism, and phylogeny: a cladistic analysis of extant African papionins using craniodental data. Journal of Human Evolution 57, 298-320.
Gilbert, C.C., Rossie, J.B., 2007. Congruence of molecules and morphology using a narrow allometric approach. Proceedings of the National Academy of Sciences 104, 11910-11914.
González-José, R., Ramírez-Rozzi, F., Sardi, M., Martínez-Abadías, N., Hernandez, M., Pucciarelli, H., 2005. Functional-cranial approach to the influence of economic strategy on skull morphology. American Journal of Physical Anthropology 128, 757-771.
Gower, J.C., 1975. Generalised Procrustes Analysis. Psychometrika 40, 33-50.
Groves, C.P., 1978. Phylogenetic and population systematics of the mangabeys (Primates: Cercopithecoidea). Primates 19, 1-34.
Groves, C.P., 2001. Primate Taxonomy. Washington, DC: Smithsonian Institution Press.
Hansen, T.F., Houle, D., 2008. Measuring and comparing evolvability and constraint in multivariate characters. Journal of Evolutionary Biology 21, 1201-1219.
Harris, E.E., 2000. Molecular systematics of the Old World monkey tribe Papionini: analysis of the total available genetic sequences. Journal of Human Evolution 38, 235-256.
Harris, E.E., Disotell, T.R., 1998. Nuclear gene trees and the phylogenetic relationships of the Mangabeys (Primates: Papionini). Molecular Biology and Evolution 15, 892-900.
Harvati, K., Weaver, T.D., 2006. Human cranial anatomy and the differential preservation of population history and climate signatures. Anatomical Record 288A, 1225-1233.
Holm, S., 1979. A simple sequential rejective multiple test procedure. Scandinavian Journal of Statistics 6, 65-70.
Holmes, M.A., Ruff, C.B., 2011. Dietary effects on development of the human mandibular corpus. American Journal Of Physical Anthropology 145, 615-628.

Jolly, C.J., 1972. The classification and natural history of *Theropithecus (Simopithecus*) (Andrews, 1916), baboons of the African Plio-Pliestocene. Bulletin of the British Museum (Natural History) Geology 22, 1-122.
Jolly, C.J., 1993. Species, subspecies and baboon systematics. In: Kimbel, W.H., Martin, L., editors. Species, species concepts and primate evolution. New York: Wiley. pp. 67-107.
Jolly, C.J., 2001. A proper study for mankind: analogies from the papionin monkeys and their implications for human evolution. Yearbook of Physical Anthropology 44, 177-204.
Jungers, W.L., Falsetti, A.B., Wall, C.E., 1995. Shape, relative size, and size-adjustments in morphometrics. Yearbook of Physical Anthropology 38, 137-161.
Klingenberg, C.P., 2011. MorphoJ: an integrated software package for geometric morphometrics. Molecular Ecology Resources 11, 353-357.
Lande, R., 1976. Natural selection and random genetic drift in phenotypic evolution. Evolution 30, 314-334.
Lande, R., Arnold, S.J., 1983. The measurement of selection on correlated characters. Evolution 37, 1210-1226.
Larsen, C., 1997. Bioarchaeology: Interpreting Behavior from the Human Skeleton. Cambridge: Cambridge University Press.
Lieberman, D.E., 1995. Testing hypotheses about recent human evolution from skulls: integrating morphology, function, development, and phylogeny. Current Anthropology 36, 159-197.
Lieberman, D.E., 2008. Speculations about the selective basis for the modern human craniofacial form. Evolutionary Anthropology 17, 55-68.
Lieberman, D.E., 2011. The Evolution of the Human Head. Cambridge, MA: Harvard University Press.
Lieberman, D.E., Pearson, O.M., Mowbray, K.M., 2000a. Basicranial influence on overall cranial shape. Journal of Human Evolution 38, 291-315.
Lieberman, D.E., Ross, C.F., Ravosa, M.J., 2000b. The primate cranial base: ontogeny, function, and integration. Yearbook of Physical Anthropology 43, 117-169.
Lieberman, D.E., Wood, B.A., Pilbeam, D.R., 1996. Homoplasy and early *Homo*: an analysis of the evolutionary relationships of *H. habilis sensu stricto* and *H. rudolfensis*. Journal of Human Evolution 30, 97-120.
Lockwood, C.A., Fleagle, J.G., 1999. The recognition and evaluation of homoplasy in primate and human evolution. Yearbook of Physical Anthropology 42, 182-232.
Lockwood, C.A., Kimbel, W.H., Lynch, J.M., 2004. Morphometrics and hominoid phylogeny: support for the chimpanzee-human clade and differentiation among great ape subspecie. Proceedings of the National Academy of Sciences 101, 4356-4360.
Lycett, S.J., Collard, M., 2005. Do homoiologies impede phylogenetic analyses of the fossil hominids? An assessment based on extant papionin craniodental morphology. Journal of Human Evolution 49, 618-642.
Lynch, M., 1990. The rate of moprhological evolution in mammals from the standpoint of the neutral expectation. American Naturalist 136, 727-741.
Mantel, N.A., 1967. The detection of disease clustering and a generalized regression approach. Cancer Research 27, 209-220.
Marroig, G., Cheverud, J.M., 2004. Did natural selection or genetic drift produce the cranial diversification of neotropical monkeys? American Naturalist 163, 417-428.
McGoogan, K., Kivell, T., Hutchison, M., Young, H., Blanchard, S., Keeth, M., Lehman, S.M., 2007. Phylogenetic diversity and the conservation biogeography of African primates. Journal of Biogeography 34, 1962-1974.
McGraw, W.W., Fleagle, J.G., 2006. Biogeography and evolution of the *Cercocebus-Mandrillus* clade: evidence from the face. In: Lehman, S., Fleagle, J.G., editors. Primate Biogeography: Progress and Prospects. New York: Springer. pp. 201-224.

Nevell, L., Wood, B., 2008. Cranial base evolution within the hominin clade. Journal of Anatomy 212, 455-468.
Newman, T.K., Jolly, C.J., Rogers, J., 2004. Mitochondrial phylogeny and systematics of baboons (*Papio*). American Journal of Physical Anthropology 124, 17-27.
Nicholson, E., Harvati, K., 2006. Quantitative analysis of human mandibular shape using three-dimensional geometric morphometrics. American Journal of Physical Anthropology 131, 368-383.
Olson, E.C., Miller, R.L., 1958. Morphological Integration. Chicago: University of Chicago Press.
Olson, T.R., 1981. Basicranial morphology of the extant hominoids and Pliocene hominids: the new material from the Hadar Formations, Ethiopia and its significance in early human evolution and taxonomy. In: Stringer, C.B., editor. Aspects of Human Evolution. London: Taylor and Francis. pp. 99-128.
Paschetta, C., de Azevedo, S., Castillo, L., Martínez-Abadías, N., Hernández, M., Lieberman, D.E., González-José, R., 2010. The influence of masticatory loading on craniofacial morphology: a test case across technological transitions in the Ohio valley. American Journal of Physical Anthropology 141, 297-314.
Perelman, P., Johnson, W.E., Roos, C., Seuánez, H.N., Horvath, J.E., Moreira, M.A.M., Kessing, B., Pontius, J., Roelke, M., Rumpler, Y. et al. , 2011. A molecular phylogeny of living primates. PLoS Genetics 7, e1001342.
Pinhasi, R., Eshed, V., Shaw, P., 2008. Evolutionary changes in the masticatory complex following the transition to farming in the southern Levant. American Journal Of Physical Anthropology 135, 136-148.
Rolian, C.C., 2009. Integration and evolvability in primate hands and feet. Evolutionary Biology 36, 100-111.
Roseman, C.C., 2004. Detecting interregionally diversifying natural selection on modern human cranial form by using matched molecular and morphometric data. Proceedings of the National Academy of Sciences 101, 12824-12829.
Roseman, C.C., Willmore, K.E., Rogers, J., Hildebolt, C., Sadler, B.E., Richtsmeier, J.T., Cheverud, J.M., 2010. Genetic and environmental contributions to variation in baboon cranial morphology. American Journal of Physical Anthropology 143, 1-12.
Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4, 406-425.
Sardi, M., Novellino, P.S., Pucciarelli, H.M., 2006. Cranio-facial morphology in the Argentine center-west: Consequences of the transition to food production. American Journal Of Physical Anthropology 130, 333-343.
Singleton, M., 2002. Patterns of cranial shape variation in the Papionini (Primates: Cercopithecinae). Journal Of Human Evolution 42, 547-578.
Smith, H.F., 2009. Which cranial regions reflect molecular distances reliably in humans? Evidence from three-dimensional morphology. American Journal of Human Biology 21, 36-47.
Smith, H.F., Ritzman, T., Otárola-Castillo, E., Terhune, C.E., 2013. A 3-D geometric morphometric study of intraspecific variation in the ontogeny of the temporal bone in modern Homo sapiens. Journal of Human Evolution 65, 479-489.
Smith, H.F., Terhune, C.E., Lockwood, C.A., 2007. Genetic, geographic, and environmental correlates of human temporal bone variation. American Journal of Physical Anthropology 134, 312-322.
Springer, M.S., Meredith, R.W., Gatesy, J., Emerling, C.A., Park, J., Rabosky, D.L., Stadler, T., Steiner, C., Ryder, O.A., Janečka, J.E. et al. , 2012. Macroevolutionary dynamics and historical biogeography of primate diversification inferred from a species supermatrix. PLoS One 7, e49521.
Strait, D.S., 2001. Integration, phylogeny, and the hominid cranial base. American Journal of Physical Anthropology 114, 273-297.

Strait, D.S., Grine, F.E., Moniz, M.A., 1997. A reappraisal of early hominid phylogeny. Journal of Human Evolution 32, 17-82.
Strasser, E., Delson, E., 1987. Cladistic analysis of cercopithecoid relationships. Journal of Human Evolution 16, 81-99.
Szalay, F.S., Delson, E., 1979. Evolutionary History of the Primates. New York: Academic Press.
Thorington, R.W., Groves, C.P., 1970. An annotated classification of the cercopithecoidea. In: Napier, J.R., Napier, P.H., editors. Old world monkeys. New York: Academic Press. pp. 629-647.
Tosi, A.J., Disotell, T.R., Morales, J.C., Melnick, D.J., 2003. Cercopithecine Y-chromosome data provide a test of competing morphological evolutionary hypotheses. Molecular Phylogenetics and Evolution 27, 510-521.
Tosi, A.J., Morales, J.C., Melnick, D.J., 1999. Y-chromosome phylogeny of the macaques (Cercopithecidae: *Macaca*). American Journal of Physical Anthropology 28S.
Vinyard, C.J., Wall, C.E., Williams, S.H., Hylander, W.L., 2003. Comparative functional analysis of skull morphology of tree-gouging primates. American Journal of Physical Anthropology 120, 153-170.
von Cramon-Taubadel, N., 2009a. Congruence of individual cranial bone morphology and neutral molecular affinity patterns in modern humans. American Journal of Physical Anthropology 140, 205-215.
von Cramon-Taubadel, N., 2009b. Revisiting the homoiology hypothesis: the impact of phenotypic plasticity on the reconstruction of human population history from craniometric data. Journal of Human Evolution 57, 179-190.
von Cramon-Taubadel, N., 2011a. Global human mandibular variation reflects differences in agricultural and hunteregatherer subsistence strategies. Proceedings of the National Academy of Sciences 108, 19546-19551.
von Cramon-Taubadel, N., 2011b. The relative efficacy of functional and developmental cranial modules for reconstructing global human population history. American Journal of Physical Anthropology 146, 83-93.
von Cramon-Taubadel, N., 2014. The microevolution of modern human cranial variation: Implications for human and primate evolution. Annals of Human Biology.
von Cramon-Taubadel, N., Smith, H.F., 2012. The relative efficacy of cranial modules for reconstructing hominoid genetic relationships: implications for the reconstruction of hominin phylogeny. Journal of Human Evolution 640-653.
Wall, C.E., 1999. A model of temporomandibular joint function in anthropoid primates based on condylar movements during mastication. American Journal of Physical Anthropology 109, 67-88.
Willmore, K.E., Roseman, C.C., Rogers, J., Richtsmeier, J.T., Cheverud, J.M., 2009. Genetic variation in baboon craniofacial sexual dimorphism. Evolution 63, 799-806.
Wood, B.A., Lieberman, D.E., 2001. Craniodental variation in *Paranthropus boisei*: A developmental and functional perspective. American Journal Of Physical Anthropology 116, 13-25.
Young, N.M., Wagner, G.P., Hallgrimsson, B., 2010. Development and the evolvability of human limbs. Proceedings of the National Academy of Sciences USA 107, 3400-3405.

Table 1. Anatomical definitions of craniomandibular landmarks digitized in the present study, and the cranial regions to which these landmarks contribute. Landmarks and cranial regions are depicted in Figures 1-4.

	Landmark	Anatomical definition	Cranial Regions
Midline			
1	Alveolon	The intersection of the interpalatal suture and a line tangent to the posterior margins of the alveolar processes	F, Pl
2	Bregma	The point where the coronal and sagittal sutures intersect	V, Fr, P
3	Basion	The point where the anterior margin of the foramen magnum intersects the midsagittal plane	C, O
4	Glabella	Most anterior midline point on the frontal bone	F, V, Fr, UF
5	Hormion	The midline point of attachment of the vomer and sphenoid bones	C, S
6	Inion	The midline point where the superior nuchal lines merge in the external occipital protuberance	C, V, O
7	Incisivon	The most posterior inferior point on the incisive fossa	F, M, Pl
8	Lambda	The midline point where the sagittal and lambdoid sutures intersect.	V, O, P
9	Nasion	The point of intersection of the nasofrontal suture and the midsagittal plane.	F, Fr, UF
10	Nasal depth	The deepest point of inflection of the nasal profile	F, UF
11	Opisthion	The point where the posterior margin of the foramen magnum intersects the midsagittal plane	C, O,C
12	Ophryon	The midline point of inflection posterior to the brow ridges	V, Fr, V
13	Palatomaxillare	The midline point of intersection of the palatine and the maxillary bones	F, M, Pl
14	Prosthion	The most anterior midline point on the maxillary alveolar process between the two central incisors	F, M, Pl
15	Sphenobasion	The midline point on the sphenooccipital suture	C, O, S
16	Subspinale	The midline point at which the inferior edge of the nasal spine becomes the anterior edge of the maxilla	F, M, Pl, UF
Bilateral			
17	Alare	The most lateral point on the nasal aperture taken perpendicular to the nasal height	F, M, UF
18	Alveolare	The most anterior point on the alveolus of the first molar	F, M, Pl
19	C/P3	The most inferior point on the external surface of the maxilla between the canine and P3	F, M, Pl
20	Posterior M2	The point on the lateral alveolus distal to M2	F, M, Pl
21	Asterion	The point where the lambdoid, parietomastoid and occipitomastoid sutures meet	V, C, O, P, T
22	Carotid canal (lat)	The most lateral point on the carotid canal	C, T
23	Carotid canal (med)	The most medial point on the carotid canal	C, T
24	Coronale	The most lateral point on the coronal suture	V, Fr, P
25	Dacryon	The point of intersection of the frontolacrimal and lacrimomaxillary suture	F, Fr, M, UF
26	Ext aud meatus (ant)	The most anterior point on the margin of the external auditory meatus	V, T
27	Ext aud meatus (pos)	The most posterior point on the margin of the external auditory meatus	V, T
28	Ext palate length	The point on the inferior surface of the maxilla that denotes the most posterior point of the alveolar process	F, M, Pl
29	Euryon (parietal)	The most lateral point on the parietals that defines the greatest cranial breadth on the parietal	V, P
30	Frontomalare orbitale	The point where the zygomaticofrontal suture crosses the orbital margin	F, Fr, Z, UF, Zt
31	Frontomalare temporale	The most lateral point on the zygomaticofrontal suture	F, Fr, Z, Zt
32	Foramen ovale (ant)	The most anterior point on the foramen ovale	V, S
33	Foramen ovale (pos)	The most posterior point on the foramen ovale	V, S

34	Foramen magnum (lat)	The most lateral point on the margin of the foramen magnum and posterior to occipital condyle	C, O
35	Frontozygomatico-sphenoid (FRED)	The point of intersection of the frontozygomatic, zygomaticosphenoid and sphenofrontal sutures	V, F, Fr, S, Z, Zt
36	Infranasion	The point of intersection of the nasofrontal, nasomaxillary and maxillofrontal sutures	F, Fr, M, UF
37	Infratemporale	The most medial point on the infratemporal crests	V, S
38	Jugular (lat)	The most inferior, lateral point on the margin of the jugular foramen	C, O, T
39	Jugular (med)	The most inferior, medial point on the margin of the jugular foramen	C, O, T
40	Jugale	The point in the depth of the notch between the temporal and frontal process of the zygomatic bone	F, Z, Zt
41	Krotaphion	The most posterior extent of the sphenoparietal suture (pterion)	V, P, S, T
42	Mandibular fossa (lat)	The most lateral point on the mandibular fossa	V, T, Zt
43	Max maxillary curve	The point in the depth of the notch between the zygomaxillary suture and the alveolar process	F, M, Pl
44	Mastoidale	The most inferior, lateral point on the mastoid process	C, T
45	Nasomaxillare	The most inferior point on the nasomaxillary suture	F, M, UF
46	Occipitocondyle (ant)	The most anterior, inferior point on the occipital condyle	C, O
47	Occipitocondyle (lat)	The most lateral, inferior point on the occipital condyle	C, O
48	Orbitale	The most inferior midpoint on the orbital margin	F, Z, UF, Zt
49	Orbitale (sup)	The most superior midpoint of the orbital margin	F, Fr, UF
50	Palatomaxillare (lat)	The most lateral point on the palato-maxillary suture	F, M, Pl
51	Petrosal	The most anterior point of the petrous element of the temporal bone	C, T, S
52	Porion	The most superior point on the margin of the external auditory meatus	V, T
53	Radiculare	The point of maximum inflection of the zygomatic processes	V, T, Zt
54	Sphenomaxillare (sup)	The most superior, lateral point of contact between the maxilla and the lateral pterygoid plate of the sphenoid	V, F, S
55	Sphenobasion (lat)	The most lateral, inferior point on the sphenooccipital synchondrosis	C, O, S
56	Sphenion	The most anterior extent of the sphenoparietal suture (Pterion)	V, Fr, P, S
57	Sphenosquamosal	The point of intersection of the infratemporal crest and sphenosquamosal suture	V, S, T, Zt
58	Stenion	The most medial point on the sphenosquamosal sutures (same as mfm)	V, C, S, T, Zt
59	Styloid foramen	The most anterior, inferior point on the styloid foramen	C, T
60	Sphenozygomatic (pos)	The most posterior, inferior point on the sphenozygomatic suture	V, F, S, Zt
61	Temporal fossa (pos)	The most posterior, inferior point on the temporal fossa	V, T, Zt
62	Zygotemporale (inf)	The most inferior point on the zygomaticotemporal suture	V, F, T, Z, Zt
63	Zygotemporale (sup)	The most superior point on the zygomaticotemporal suture	V, F, T, Z, Zt
64	Zygomaxillare	The most inferior, anterior point on the zygomaticomaxillary suture	F, M, Z, UF, Zt
65	Zygoorbitale	The point where the zygomaticomaxillary suture intersects with the inferior orbital margin	F, M, Z, UF, Zt
66	Zygion	The most lateral point on the surface of the zygomatic arch	V, Z, Zt
Mandibular			
1	Condylion (med)	The most medial point on the superior surface of the mandibular condyle	Mn
2	Condylion (lat)	The most lateral point on the superior surface of the mandibular condyle	Mn
3	I2 (lat)	The most distal point on the alveolus lateral to I2	Mn
4	Canine/P3 (lat)	The most lateral point on the alveolus between the canine and P3	Mn

5	Alveolare	The most lateral point on the alveolus anterior to the M1	Mn
6	M1 (pos)	The most lateral point on the alveolus posterior to the M1	Mn
7	M2 (pos)	The most lateral point on the alveolus posterior to the M2	Mn
8	Gonion	The point of maximum curvature on the posterior-inferior border where posterior ramus and corpus intersect	Mn
9	Sigmoid notch	The most superior point of maximum inflection in the depth of the sigmoid notch	Mn
10	Coronion	The most superior point on the coronoid process	Mn
11	Infradentale	The most superior midline point on the buccal surface of the alveolus	Mn
12	Pogonion	The most anterior midline point on the mental eminence	Mn
13	Gnathion	The most inferior midline point on the mandibular symphysis	Mn
14	Mandibular orale	The most superior midline point on the lingual surface of the alveolus	Mn

V = vault, F = face, C = chondrocranium, Fr = frontal, O = occipital, P = parietal, T = temporal, M = maxilla, Mn= mandible, S = sphenoid, Z = zygomatic, Pl = palatomaxilla, Zt = zygotemporal, UF = upper face. lat = lateral, med = medial, ant = anterior, pos = posterior, sup = superior, inf = inferior.

Table 2. Papionin taxa included in the present study, sample sizes, and the museum collections from which each was derived.

Taxa		**Sample sizes**		
	Museum Collections	Males	Females	Total
Cercocebus agilis	AMNH, MCZ	11	10	21
Cercocebus atys	FMNH, MCZ, NMNH	8	8	16
Cercocebus torquatus	AMNH, FMNH, MCZ, NMNH	18	10	28
Lophocebus albigena	AMNH, FMNH, NMNH	18	12	30
Macaca fascicularis	NMNH	17	13	30
*Macaca mulatta**	MCZ, NMNH	6	22	28
Macaca nemestrina	AMNH, FMNH, MCZ, NMNH	18	22	40
Macaca sylvanus	AMNH, MCZ, NMNH	15	11	26
Mandrillus sphinx	AMNH, FMNH, MCZ, NMNH	20	11	31
Papio anubis	AMNH, FMNH, NMNH	16	13	29
*Papio cynocephalus**	AMNH, FMNH, MCZ, NMNH	13	7	20
*Papio hamadryas**	AMNH, FMNH, MCZ, NMNH	17	4	21
*Papio papio**	AMNH, FMNH, NMNH	15	5	20
Papio ursinus	FMNH, NMNH	20	10	30

AMNH = American Museum of Natural History (New York), FMNH = Field Museum of Natural History (Chicago), MCZ = Museum of Comparative Zoology (Harvard), NMNH = National Museum of Natural History (Washington DC).

*Taxon included in combined-sex analyses only.

Table 3. Results of Mantel tests between morphological and molecular matrices for each of the four data sets: complete 14 taxon combined-sex data set, reduced 10 taxon combined-sex data set, reduced 10 taxon male-only data set, and reduced 10 taxon female-only data set. Cranial regions are ordered according to *r*-value in the combined-sex 14 taxon analysis from strongest (largest *r*-values) to the weakest correlations. Significant correlations ($\alpha < 0.05$) are indicated in bold. Correlations that remain significant after a sequential Bonferroni correction are indicated with an asterisk. Func-Dev = Functional-Developmental regions.

Cranial region category	**Prediction***	**Cranial region**	**14 taxon Mixed Sex**		**10 taxon Mixed Sex**		**10 taxon Males**		**10 taxon Females**	
			r-value	*p-value*	*r-value*	*p-value*	*r-value*	*p-value*	*r-value*	*p-value*
Entire Cranium			**0.477***	**0.001**	0.066	0.319	0.096	0.230	0.205	0.074
Func-Dev	3	Chondrocranium	**0.557***	**<0.001**	**0.440***	**0.006**	**0.513***	**0.004**	-0.020	0.457
	1	Splanchnocranium	**0.413***	**0.002**	0.205	0.091	0.170	0.122	0.149	0.154
	2	Neurocranium	**0.335***	**<0.001**	**0.290**	**0.032**	0.074	0.313	0.100	0.222
Bones	1	Zygomatic	**0.470***	**<0.001**	**0.333**	**0.016**	**0.272**	**0.040**	0.249	0.056
	2	Maxilla	**0.453***	**<0.001**	0.100	0.250	0.160	0.124	0.079	0.268
	3	Temporal	**0.452***	**<0.001**	**0.348**	**0.012**	**0.362***	**0.016**	0.223	0.055
	3	Occipital	**0.419***	**<0.001**	**0.314**	**0.017**	0.102	0.225	**0.284**	**0.023**
	2	Frontal	**0.376***	**0.007**	0.204	0.093	0.154	0.122	0.185	0.110
	2	Parietal	0.175	0.088	0.082	0.274	0.019	0.421	0.173	0.120
	2	Sphenoid	0.048	0.366	-0.091	0.274	-0.031	0.419	0.107	0.217
Masticatory	2	Chondrocranium	**0.557***	**<0.001**	**0.440**	**0.006**	**0.513**	**0.004**	-0.020	0.457
	1	Zygotemporal	**0.498***	**0.001**	**0.325**	**0.022**	0.233	0.064	**0.281**	**0.037**
	3	Palatomaxilla	**0.440***	**0.001**	0.206	0.073	0.119	0.188	0.163	0.139
	1	Upper face	**0.425***	**0.002**	0.160	0.158	0.163	0.133	0.114	0.194
	3	Mandible	**0.418***	**<0.001**	**0.237**	**0.043**	0.220	0.079	0.082	0.297
	2	Neurocranium	**0.335***	**<0.001**	**0.290**	**0.032**	0.074	0.313	0.100	0.222

***** Numbers signify which regions in each subset (functional/developmental, masticatory strain, and individual bones) were found to be significantly more genetically congruent in hominoids based on Dow-Cheverud tests (see von Cramon-Taubadel and Smith, 2012). Thus, 1 is significantly better than 2, which is significantly better than 3.

Table 4. Results of all Dow-Cheverud test comparisons for the 14 taxon combined-sex data set. Cranial regions are ordered according to the results of the Mantel tests from strongest (largest *r*-values) to the weakest correlations. Lower diagonals = p1Z values, upper diagonals = *p*-values. Significant differences ($\alpha < 0.05$) are indicated in bold. Differences that remain significant after a sequential Bonferroni correction are indicated with an asterisk. Chondro = Chondrocranium, Neuro= Neurocranium, Palatomax = Palatomaxilla, Zygotemp = Zygotemporal, Zygomat= Zygomatic.

Func-Dev	Chondro	Cranium	Splanchno	Neuro				
Chondro		0.086	**0.028**	**0.015**				
Cranium	0.150		0.132	0.057				
Splanchno	-0.218	-0.214		0.175				
Neuro	-0.271	-0.194	0.111					
Bones	Cranium	Zygomat	Maxilla	Temporal	Occipital	Frontal	Parietal	Sphenoid
Cranium		0.480	0.320	0.365	0.292	0.079	**0.013**	**<0.001***
Zygomatic	-0.010		0.413	0.403	0.344	0.109	**0.030**	**0.003**
Maxilla	-0.045	0.026		0.474	0.398	0.108	**0.020**	**0.002***
Temporal	-0.036	-0.022	-0.001		0.385	0.269	**0.016**	**0.001***
Occipital	-0.067	0.054	-0.036	0.041		0.386	**0.027**	**<0.001***
Frontal	-0.168	0.129	0.137	0.075	0.044		0.064	**0.005**
Parietal	-0.284	0.238	-0.268	0.267	-0.231	-0.191		0.164
Sphenoid	-0.432	0.350	-0.396	0.401	-0.381	-0.331	-0.139	
Masticatory	Chondro	Zygotemp	Cranium	Palatomax	Upper face	Mandible	Neuro	
Chondro		0.194	0.086	0.100	**0.035**	0.060	**0.015**	
Zygotemp	-0.096		0.362	0.224	0.133	0.151	**0.020**	
Cranium	0.150	0.043		0.316	0.198	0.332	0.057	
Palatomax	-0.152	0.086	-0.058		0.355	0.339	0.141	
Upper face	-0.196	-0.123	-0.087	-0.033		0.455	0.150	
Mandible	-0.186	0.121	-0.045	0.045	0.011		0.191	
Neuro	-0.271	0.229	-0.194	0.126	0.114	-0.108		

Table 5. Results of all Dow-Cheverud test comparisons for the reduced 10 taxon combined-sex data set. Cranial regions are ordered according to the results of the Mantel tests for this data set from strongest (largest *r*-values) to the weakest correlations. Lower diagonals = p1Z values, upper diagonals = *p*-values. Significant differences ($\alpha < 0.05$) are indicated in bold. Differences that remain significant after a sequential Bonferroni correction are indicated with an asterisk. Chondro = Chondrocranium, Neuro= Neurocranium, Palatomax = Palatomaxilla, Zygotemp = Zygotemporal, Zygomat= Zygomatic.

Func-Dev	Chondro	Neuro	Splanchno	Cranium				
Chondro		0.095	**0.024**	**0.010**				
Neuro	-0.183		0.192	0.054				
Splanchno	-0.301	-0.124		0.136				
Cranium	-0.333	-0.231	-0.166					
Bones	Temporal	Zygomat	Occipital	Frontal	Maxilla	Parietal	Cranium	Sphenoid
Temporal		0.442	0.380	0.138	**0.012**	**0.022**	**0.006**	**0.002***
Zygomat	-0.190		0.467	0.097	**0.019**	0.071	**0.046**	**0.011**
Occipital	-0.038	-0.169		0.210	0.054	**0.017**	0.052	**<0.001***
Frontal	-0.159	-0.179	-0.111		0.135	0.195	0.166	**0.022**
Maxilla	-0.326	-0.303	-0.213	-0.171		0.470	0.407	0.066
Parietal	-0.261	-0.209	-0.286	-0.126	-0.019		0.455	0.058
Cranium	-0.352	-0.239	-0.227	-0.130	-0.032	-0.014		0.101
Sphenoid	-0.421	-0.338	-0.447	-0.287	-0.215	-0.217	-0.168	
Masticatory	Chondro	Zygotemp	Vault	Mandible	Palatomax	Upper face	Cranium	
Chondro		0.139	0.095	**0.013**	**0.022**	**0.004**	**0.010**	
Zygotemp	-0.159		0.344	0.182	0.114	**0.042**	**0.010**	
Vault	-0.183	-0.056		0.334	0.254	0.128	**0.050**	
Mandible	-0.331	-0.124	-0.060		0.350	0.173	0.133	
Palatomax	-0.287	-0.147	-0.092	-0.005		0.253	0.203	
Upper face	-0.371	-0.246	-0.167	-0.123	-0.100		0.252	
Cranium	-0.333	-0.312	-0.231	-0.160	-0.129	-0.098		

Table 6. Results (r-values) of the allometrically size corrected Mantel tests compared with the original Mantel test results. Cranial regions are ordered according to their original Mantel test results in the 14 taxon data set. All significant correlations ($\alpha < 0.05$) are highlighted in bold. Correlations that remain significant after a sequential Bonferroni correction are indicated with an asterisk. Zygotemp = Zygotemporal, Chondro = Chondrocranium, Neuro= Neurocranium, Palatomax = Palatomaxilla.

Cranial Regions	**14 taxon Mixed Sex**				**10 taxon Mixed Sex**				**10 taxon Male-only**				**10 taxon Female-only**			
	Original r-value	*Original p-value*	*New r-value*	*New p-value*	*Original r-value*	*Original p-value*	*New r-value*	*New p-value*	*Original r-value*	*Original p-value*	*New r-value*	*New p-value*	*Original r-value*	*Original p-value*	*New r-value*	*New p-value*
Cranium	**0.477***	**0.001**	**0.304***	**0.012**	0.066	0.319	0.135	0.190	0.096	0.230	**0.235**	**0.043**	0.205	0.074	**0.354**	**0.006**
Chondro	**0.557***	**<0.001**	**0.512***	**0.001**	**0.440***	**0.006**	**0.446***	**0.002**	**0.513***	**0.004**	**0.453***	**0.001**	-0.020	0.457	0.089	0.275
Splanchno	**0.413***	**0.002**	**0.387***	**<0.001**	0.205	0.091	**0.514***	**0.002**	0.170	0.122	**0.388***	**0.002**	0.149	0.154	**0.440***	**0.001**
Neuro	**0.335***	**<0.001**	**0.351***	**0.004**	**0.290**	**0.032**	**0.334***	**0.028**	0.074	0.313	**0.235***	**0.047**	0.100	0.222	0.050	0.361
Zygomatic	**0.470***	**<0.001**	**0.487***	**0.001**	**0.333**	**0.016**	**0.348**	**0.019**	**0.272**	**0.040**	**0.299**	**0.026**	0.249	0.056	**0.268**	**0.047**
Maxilla	**0.453***	**<0.001**	**0.385***	**0.002**	0.100	0.250	0.141	0.163	0.160	0.124	**0.260**	**0.031**	0.079	0.268	0.011	0.419
Temporal	**0.452***	**<0.001**	**0.330***	**<0.001**	**0.348**	**0.012**	**0.416***	**0.002**	**0.362***	**0.016**	**0.431***	**0.001**	0.223	0.055	**0.241**	**0.042**
Occipital	**0.419***	**<0.001**	**0.432***	**0.001**	**0.314**	**0.017**	**0.324***	**0.009**	0.102	0.225	0.200	0.087	**0.284**	**0.023**	**0.271**	**0.028**
Frontal	**0.376***	**0.007**	**0.348***	**0.002**	0.204	0.093	0.202	0.084	0.154	0.122	0.164	0.129	0.185	0.110	0.222	0.077
Parietal	0.175	0.088	0.142	0.097	0.082	0.274	0.096	0.261	0.019	0.421	0.080	0.283	0.173	0.120	0.079	0.253
Sphenoid	0.048	0.366	0.034	0.349	-0.091	0.274	-0.118	0.207	-0.031	0.419	-0.036	0.441	0.107	0.217	0.072	0.293
Zygotemp	**0.498***	**0.001**	**0.356***	**0.001**	**0.325**	**0.022**	**0.362***	**0.004**	0.233	0.064	**0.408***	**0.003**	**0.281**	**0.037**	0.199	0.069
Palatomax	**0.440***	**0.001**	**0.360***	**<0.001**	0.206	0.073	**0.398***	**<0.001**	0.119	0.188	**0.449***	**0.004**	0.163	0.139	0.199	0.090
Upper face	**0.425***	**0.002**	**0.420***	**<0.001**	0.160	0.158	**0.347***	**<0.001**	0.163	0.133	**0.388***	**0.001**	0.114	0.194	**0.248**	**0.037**
Mandible	**0.418***	**<0.001**	**0.317***	**0.001**	**0.237**	**0.043**	**0.318***	**0.006**	0.220	0.079	**0.412***	**0.001**	0.082	0.297	0.071	0.385

Table 7. Results of all Dow-Cheverud test comparisons for the size-adjusted full 14 taxon combined-sex data set. Cranial regions are ordered according to the results of the Mantel tests for this data set from strongest (largest *r*-values) to the weakest correlations. Lower diagonals = p1Z values, upper diagonals = *p*-values. Significant differences ($\alpha < 0.05$) are indicated in bold. Differences that remain significant after a sequential Bonferroni correction are indicated by an asterisk. Chondro = Chondrocranium, Neuro= Neurocranium, Palatomax = Palatomaxilla, Zygotemp = Zygotemporal, Zygomat= Zygomatic.

Func-Dev	Chondro	Splanchno	Neuro	Cranium				
Chondro		0.190	0.119	**0.036**				
Splanchno	-0.122		0.390	0.202				
Neuro	-0.152	-0.035		0.359				
Cranium	0.228	0.121	0.044					
Bones	Zygomat	Occipital	Maxilla	Frontal	Temporal	Cranium	Parietal	Sphenoid
Zygomat		0.336	0.234	0.067	0.115	0.110	**0.010**	**0.003**
Occipital	0.050		0.368	0.264	0.194	0.197	**0.013**	**<0.001***
Maxilla	0.094	0.048		0.392	0.329	0.219	**0.045**	**0.017**
Frontal	0.174	0.082	0.037		0.449	0.421	0.061	**0.006**
Temporal	0.132	-0.112	-0.067	-0.014		0.408	0.105	**0.022**
Cranium	0.159	0.125	0.116	0.038	0.035		0.138	0.078
Parietal	0.267	-0.269	-0.229	-0.194	0.176	-0.144		0.198
Sphenoid	-0.377	-0.391	-0.288	-0.312	-0.259	-0.215	-0.111	
Masticatory	Chondro	Upper Face	Palatomax	Zygotemp	Neuro	Mandible	Cranium	
Chondro		0.240	0.122	0.115	0.119	0.124	**0.036**	
Upper Face	-0.098		0.292	0.257	0.350	0.313	0.281	
Palatomax	-0.143	-0.086		0.495	0.465	0.463	0.298	
Zygotemp	-0.169	-0.091	0.005		0.484	0.454	0.278	
Neuro	-0.152	0.057	0.007	0.005		0.439	0.359	
Mandible	-0.154	0.060	-0.012	-0.184	-0.019		0.263	
Cranium	0.228	0.095	0.069	0.081	0.044	0.095		

Table 8. Results of all Dow-Cheverud test comparisons for the size-adjusted combined-sex 10 taxon data set. Cranial regions are ordered according to the results of the Mantel tests for this data set from strongest (largest *r*-values) to the weakest correlations. Lower diagonals = p1Z values, upper diagonals = *p*-values. Significant differences ($\alpha < 0.05$) are indicated in bold. Differences that remain significant after a sequential Bonferroni correction are indicated by an asterisk. Chondro = Chondrocranium, Neuro= Neurocranium, Palatomax = Palatomaxilla, Zygotemp = Zygotemporal, Zygomat= Zygomatic.

Func-Dev	Splanchno	Chondro	Neuro	Cranium				
Splanchno		0.332	0.111	**0.006***				
Chondro	0.061		0.247	**0.016**				
Neuro	0.179	-0.106		0.128				
Cranium	0.346	0.310	0.164					
Bones	Temporal	Zygomatic	Occipital	Frontal	Maxilla	Cranium	Parietal	Sphenoid
Temporal		0.386	0.244	0.107	**0.014**	**0.012**	**0.011**	**0.003**
Zygomatic	-0.053		0.462	0.120	0.125	0.116	0.094	**0.005**
Occipital	0.105	0.016		0.234	0.088	0.115	**0.024**	**0.005**
Frontal	0.175	0.180	0.105		0.338	0.312	0.258	**0.019**
Maxilla	0.301	0.179	0.192	-0.063		0.498	0.385	**0.038**
Cranium	0.328	0.174	0.181	0.006	0.006		0.391	0.056
Sphenoid	0.391	0.370	-0.356	-0.291	-0.234	-0.228		0.090
Parietal	0.322	0.177	-0.280	0.092	-0.044	-0.033	-0.191	
Masticatory	Chondro	Palatomax	Zygotemp	UpperFace	Neuro	Mandible	Cranium	
Chondro		0.389	0.254	0.253	0.247	0.183	**0.016**	
Palatomax	-0.039		0.375	0.497	0.387	0.235	**0.049**	
Zygotemp	-0.105	-0.036		0.480	0.426	0.413	**0.039**	
UpperFace	-0.097	-0.064	0.001		0.351	0.351	0.072	
Neuro	-0.106	0.051	0.013	0.010		0.445	0.128	
Mandible	-0.136	0.174	0.024	0.050	-0.013		0.109	
Cranium	0.310	0.232	0.236	0.207	0.164	0.174		

Table 9. Results of all Dow-Cheverud test comparisons for the size-adjusted male-only 10 taxon data set. Cranial regions are ordered according to the results of the Mantel tests for this data set from strongest (largest *r*-values) to the weakest correlations. Lower diagonals = p1Z values, upper diagonals = *p*-values. Significant differences ($\alpha < 0.05$) are indicated in bold. No differences remained significant after a sequential Bonferroni correction. Chondro = Chondrocranium, Neuro= Neurocranium, Palatomax = Palatomaxilla, Zygotemp = Zygotemporal, Zygomat= Zygomatic.

Func-Dev	Chondro	Splanchno	Cranium	Neuro				
Chondro		0.357	0.089	0.107				
Splanchno	-0.053		0.147	0.142				
Cranium	0.201	0.160		0.485				
Neuro	-0.178	0.162	0.003					
Bones	Temporal	Zygomat	Maxilla	Cranium	Occipital	Frontal	Parietal	Sphenoid
Temporal		0.235	0.094	**0.035**	0.290	**0.038**	**0.006**	**0.016**
Zygomat	0.116		0.413	0.349	0.290	0.144	0.152	**0.041**
Maxilla	0.194	0.031		0.400	0.357	0.265	0.104	0.347
Cranium	0.246	0.052	0.028		0.401	0.290	0.147	**0.037**
Occipital	-0.037	0.070	-0.055	-0.037		0.441	0.123	0.115
Frontal	0.252	0.159	0.087	-0.077	0.033		0.294	0.101
Parietal	0.309	0.154	-0.172	-0.141	-0.162	-0.082		0.242
Sphenoid	0.310	0.246	-0.055	-0.236	-0.165	-0.169	-0.097	
Masticatory	Chondro	Palatomax	Zygotemp	Upper Face	Neuro	Mandible	Cranium	
Chondro		0.517	0.371	0.323	0.107	0.395	0.089	
Palatomax	-0.002		0.394	0.308	0.094	0.393	0.093	
Zygotemp	-0.048	-0.036		0.404	0.084	0.495	0.080	
Upper Face	-0.063	-0.070	-0.028		0.188	0.378	0.123	
Neuro	-0.178	0.176	0.180	0.131		0.128	0.485	
Mandible	-0.042	0.036	-0.005	-0.042	-0.150		0.132	
Cranium	0.201	0.189	0.192	0.164	0.003	0.167		

Table 10. Results of all Dow-Cheverud test comparisons for the size-adjusted female-only 10 taxon data set. Cranial regions are ordered according to the results of the Mantel tests for this data set from strongest (largest *r*-values) to the weakest correlations. Lower diagonals = p1Z values, upper diagonals = *p*-values. Significant differences ($\alpha < 0.05$) are indicated in bold. No differences remained significant after a sequential Bonferroni correction. Chondro = Chondrocranium, Neuro= Neurocranium, Palatomax = Palatomaxilla, Zygotemp = Zygotemporal, Zygomat= Zygomatic.

Func-Dev	Splanchno	Cranium	Chondro	Neuro				
Splanchno		0.188	**0.013**	**0.019**				
Cranium	0.128		0.058	**0.031**				
Chondro	0.315	-0.208		0.437				
Neuro	-0.295	-0.265	-0.027					
Bones	Cranium	Occipital	Zygomat	Temporal	Frontal	Parietal	Sphenoid	Maxilla
Cranium		0.309	0.347	0.181	0.193	0.060	**0.050**	**0.006**
Occipital	-0.070		0.489	0.424	0.406	0.140	0.097	**0.035**
Zygomat	-0.062	-0.002		0.453	0.413	0.172	0.163	0.117
Temporal	-0.127	-0.027	0.018		0.476	0.121	0.169	**0.032**
Frontal	-0.126	0.036	0.039	0.013		0.190	0.240	0.059
Parietal	-0.209	-0.156	0.122	0.158	-0.155		0.501	0.304
Sphenoid	-0.224	-0.176	0.133	0.143	-0.104	-0.006		0.323
Maxilla	-0.352	0.235	0.179	0.270	-0.212	0.075	0.055	
Masticatory	Cranium	Upper Face	Zygotemp	Palatomax	Chondro	Mandible	Neuro	
Cranium		0.215	**0.033**	0.080	0.058	**0.019**	**0.031**	
Upper Face	-0.117		0.354	0.389	0.219	0.103	0.151	
Zygotemp	-0.258	0.051		0.494	0.240	0.095	0.173	
Palatomax	-0.213	0.049	-0.001		0.260	0.233	0.182	
Chondro	-0.208	0.123	0.094	0.082		0.462	0.437	
Mandible	-0.295	0.170	0.182	0.115	-0.014		0.446	
Neuro	-0.265	0.138	0.129	0.122	-0.027	-0.021		

Table 11. Results of all Dow-Cheverud test comparisons between the combined-sex and male-only and female-only 10 taxon size-adjusted datasets. Cranial regions are ordered according to the results of the Mantel tests from strongest (largest *r*-values) to the weakest correlations. Significant differences ($\alpha < 0.05$) are indicated in bold. Differences that remain significant after a sequential Bonferroni correction are indicated by an asterisk.

Cranial region category	**Cranial region**	**Size-adjusted males *vs.* combined sex**		**Size-adjusted females *vs.* combined sex**	
		p1z-value	*p-value*	*p1z-value*	*p-value*
Entire Cranium		0.109	0.235	**-0.229**	**0.050**
Func-Dev	Chondrocranium	0.008	0.458	**-0.390***	**0.001**
	Splanchnocranium	-0.172	0.120	-0.121	0.204
	Neurocranium	-0.131	0.203	**-0.292***	**0.017**
Bones	Zygomatic	-0.129	0.180	-0.185	0.096
	Occipital	**-0.253**	**0.038**	-0.060	0.326
	Maxilla	0.194	0.083	**-0.219**	**0.049**
	Frontal	-0.058	0.358	0.030	0.412
	Temporal	0.024	0.425	**-0.313***	**0.008**
	Parietal	-0.059	0.355	-0.031	0.394
	Sphenoid	0.182	0.088	**0.218**	**0.050**
Masticatory	Upper face	0.117	0.195	-0.187	0.094
	Palatomaxilla	0.083	0.329	**-0.369***	**0.002**
	Zygotemporal	0.128	0.171	**-0.334***	**0.003**
	Mandible	0.188	0.098	**-0.305***	**0.008**

FIGURES

Figure 1. Anatomical position of all 116 cranial landmarks digitized. Full anatomical descriptions of each landmark can be found in Table 1. Different colors here depict the three main cranial regions defined using developmental and functional criteria; pink = chondrocranium, yellow = splanchnocranium, blue = neurocranium.

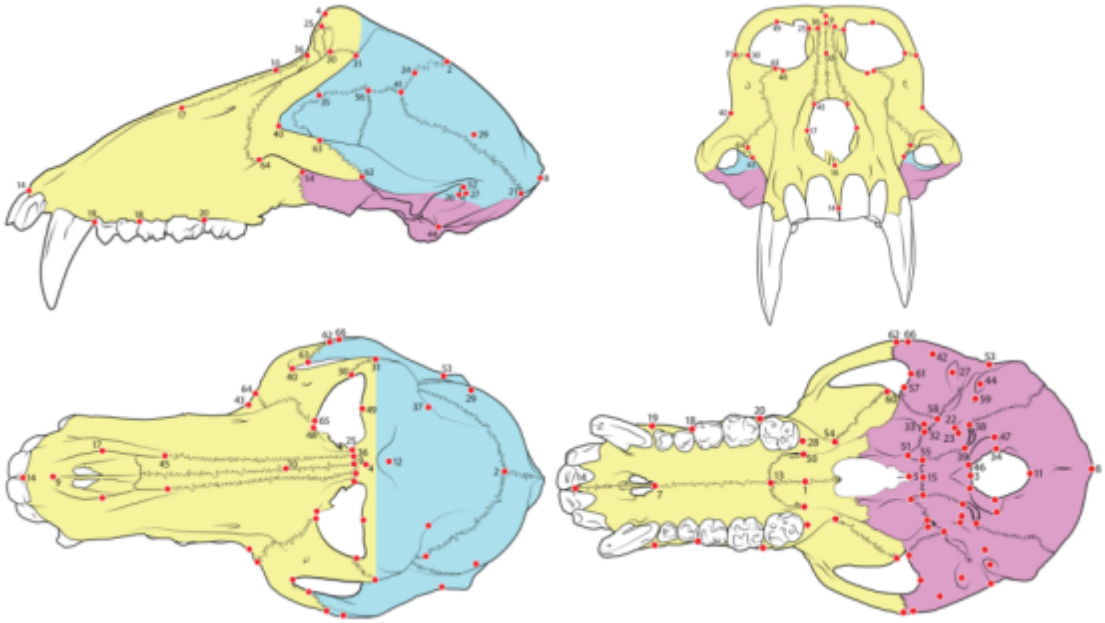

Figure 2. Landmark configurations used to describe the shape of cranial regions defined on the basis of masticatory function; yellow = upper face, blue = zygotemporal, and pink = palatomaxilla. All landmark descriptions can be found in Figure 1 and Table 1.

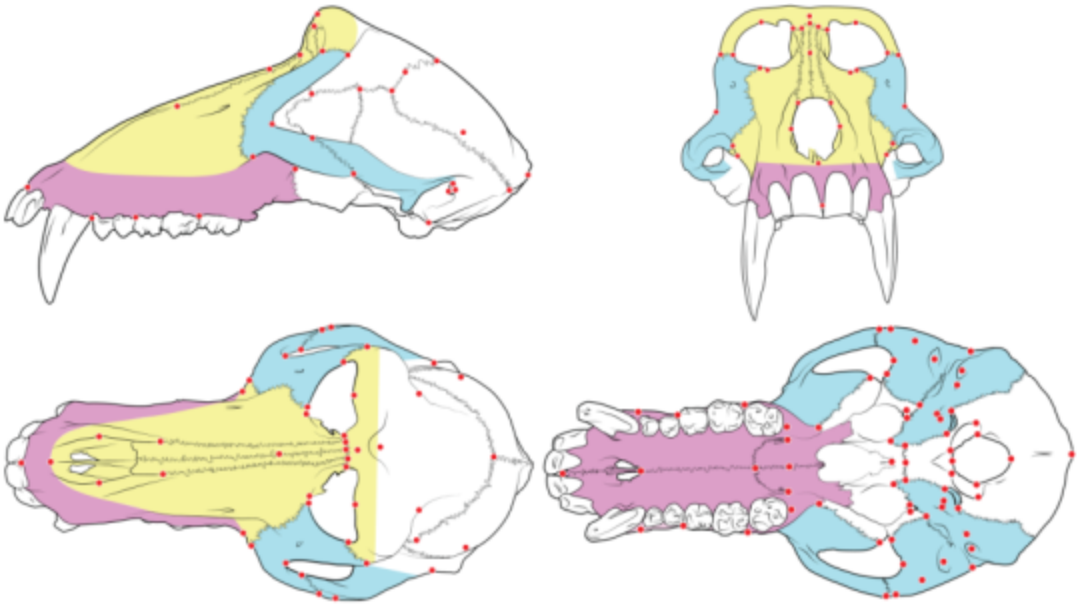

Figure 3. Landmark configurations used to describe the shape of cranial regions defined on the basis of individual cranial bones; yellow = temporal, dark blue = zygomatic, light blue = frontal, purple = parietal, orange = occipital, green = sphenoid, pink = maxilla. All landmark descriptions can be found in Figure 1 and Table 1.

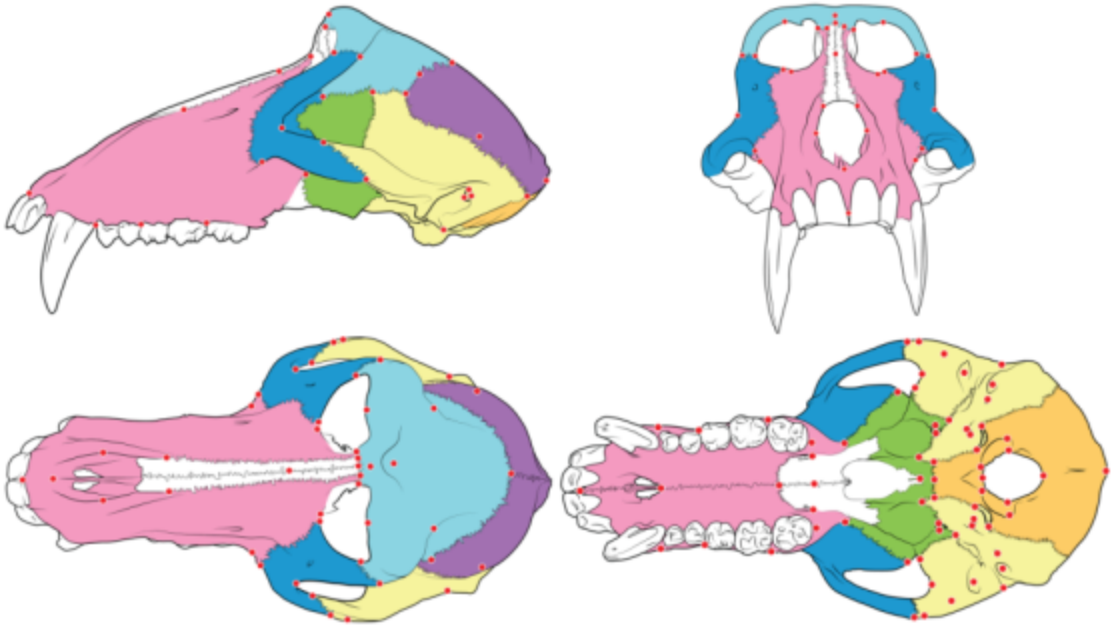

Figure 4. Landmark configuration used to describe the shape of the mandible. Numbers match the anatomical descriptions given in Table 1.

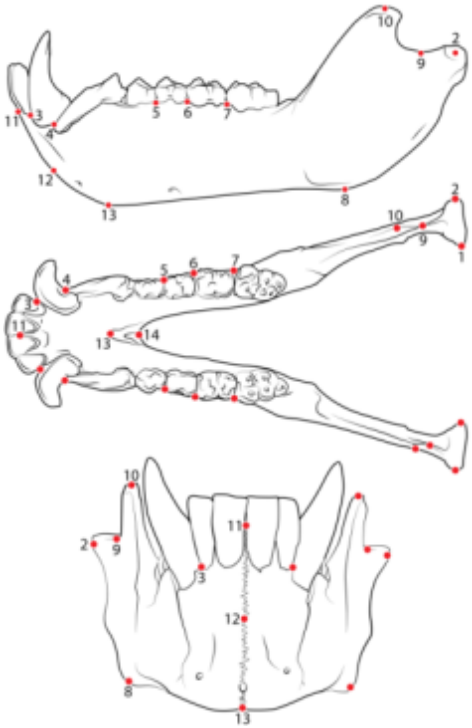

Figure 5. Consensus molecular phylogeny for the papionin taxa included in the present study based on the current published literature. Nodes are numbered (italics) according to Perelman *et al*. (2011) and branch lengths are taken from the maximum likelihood values presented there. Branch lengths in bold are derived from additional published sources on specific molecular relationships not covered by Perelman *et al*. (2011); Newman *et al*. (2004) and McGoogan *et al*. (2007) in the case of *Papio cynocephalus* and *Papio ursinus,* and McGoogan *et al*. (2007) and Springer *et al*. (2012) in the case of *Cercocebus atys*. Please see text for further details.